Loyola University Chicago
Loyola eCommons

Master's Theses | Theses and Dissertations

1991

Structure and Function of Human Alphoid DNA

Darryl S. Schlegel
Loyola University Chicago

Follow this and additional works at: https://ecommons.luc.edu/luc_theses

 Part of the Biology Commons

Recommended Citation

Schlegel, Darryl S., "Structure and Function of Human Alphoid DNA" (1991). *Master's Theses*. 3724.
https://ecommons.luc.edu/luc_theses/3724

This Thesis is brought to you for free and open access by the Theses and Dissertations at Loyola eCommons. It has been accepted for inclusion in Master's Theses by an authorized administrator of Loyola eCommons. For more information, please contact ecommons@luc.edu.

This work is licensed under a Creative Commons Attribution-Noncommercial-No Derivative Works 3.0 License.
Copyright © 1991 Darryl S. Schlegel

STRUCTURE AND FUNCTION OF HUMAN ALPHOID DNA

by

Darryl S. Schlegel

Thesis Submitted to the Faculty of the Graduate School of Loyola University of Chicago in Partial Fulfillment of the Requirements for the Degree of

Master of Science

January

1991

Copyright, 1991, Darryl Schlegel
All rights reserved.

This thesis is dedicated to my son, Kyle, whose life began as this work neared its completion.

ABSTRACT

Human alphoid DNA consists of a group of repetitive sequences which all have a basic 170 bp repeating unit and are located at the centromeres of all chromosomes. Restriction maps for two 680 bp EcoRI human alphoid clones are presented. The maps reinforce the concept that sequence heterogeneity exists in the 680 bp fragments to a greater extent than previously believed. I have also described several stringency conditions which allow for the elimination of cross-hybridization between seven different alphoid clones and a qualitative determination of sequence relatedness between the clones. This method was implemented in a study of genomic organization and copy number for each of these alphoid sequences. A demonstrated sequence and structural relationship is shown to exist between the redefined functional yeast centromere sequence (CEN) and variants of two alphoid family sequences. Data resulting from an assay to determine the functional capability of a single alphoid is reported.

ACKNOWLEDGMENTS

The work presented in this thesis could not have begun if not for the drive and insight of Dr. Jeffery Doering. It certainly could not have been completed if not for his patience, understanding and more patience. It was an honor to work close to, and observe, the manifestation of dedication. I would also like to thank Anne E. Burket for guidance throughout this work.

VITA

The author, Darryl Scott Schlegel is the son of Robert Woodrow Schlegel and Joan (Rezab) Kochan. He was born June 3, 1958, in Ft. Lauderdale, Florida.

His elementary education was obtained in the public school system of suburban Chicago, Illinois. His secondary education was completed in 1979, with a GED received through Harper College in Palatine, Illinois. In September, 1980, Mr. Schlegel entered Harper and in September, 1982, transferred to Loyola University in Chicago, Illinois, where he was awarded a Bachelor of Science in Biology in 1985. In September, 1985, Mr. Schlegel was granted a partial assistantship in biology at Loyola aiding in his completion of the Master of Science in 1990. In September of 1987 Mr. Schlegel entered Rush Medical College in Chicago, where he is expected to receive his M.D. degree in the spring of 1991.

PUBLICATIONS

Doering, J.L., Burket, A.E., Hanlon, K.M., and Schlegel, D.S. (1986). New Subfamilies of Human Alphoid Repetitive DNA. J. Cell Biol. 103:491a.

Doering J.L., Palamidis, E., and Schlegel, D.S. (1987). Genomic Organization of Human Alphoid Repetitive Sequence Families. J. Cell Biol. 105:241a.

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

LIST OF FIGURES

Figure Page

INTRODUCTION

Contained within the genomes of most eucaryotes are repetitive DNA sequences (Brutlag et al., 1980). The best estimate is that at least 20% of the human genome consists of repetitive DNA (Schmid et al., 1982; Sharp, 1983); over 1% of the genome can be accounted for by the EcoRI family alone (Furlong et al., 1986). Families of repetitive sequences can contain as many as 10^5 members.

The repeats may be in either the interspersed or tandem form. The sequence contained within both types of repeats are characteristic of the individual species, and have been detected in animal (Brutlag et al., 1980; Singer, 1982), higher plant (Walbot et al., 1979), and yeast (Clarke et al., 1986) genomes. The interspersed repeats are repetitive sequences present as single copies at many genomic loci. In many eucaryotes, such as the sea urchin and human, these interspersed repeats are short and are typically about 300 bp in length (Schmid et al., 1982; Davidson et al., 1975). Unlike the sea urchin and most other eucaryotes, where no single repeat dominates the genome, the human genome contains two predominant families

of interspersed sequences, called the Alu and Kpn families (Schmid et al., 1982). In tandemly repeated sequences, the basic unit of repetition is present as multiple copies arranged end to end at a single or small number of loci. Initially, tandemly repeated DNA was identified as "satellites",

sequences separated from the "main band" of genomic DNA through the use of isopycnic density gradients (Corneo et al., 1967, 1970, 1972; Mitchell et al., 1979). Digestion of human genomic DNA with various restriction endonucleases reveals another group of tandemly repeated sequences identified as specific sized fragments, visible above a smear of heterogenous-sized fragments when displayed on agarose gels. Within this second group are the tandem repeats termed the "alphoid family" (Maio et al., 1981). The nomenclature is due to the relatedness of these repeats to the α-satellite of the African Green Monkey (AGM) (Singer et al., 1982; Maio et al., 1981), the first species in which such repeats were discovered (Maio, 1971,; Wu & Manulelidis, 1980). The alphoid repeats are the best studied of the tandemly organized sequences. Alphoid DNA comprises the bulk of the genome in Kangaroo rat, a moderate amount (up to 24%) in a wide variety of primates, a small fraction in humans (2 to 3%), and essentially none of the genome of the yeast Saccharomyces cerevisiae (Maio,

1971; Wu & Manuelidis, 1980; Potter & Jones, 1983).

The high degree of sequence similarity among alphoid repeats in eucaryotes can be explained by the homogenization of the ancestral fundamental monomeric unit within the primitive eucaryotic cells (Wu & Manuelidis, 1980; Jorgensen et al., 1986). A high rate of homogenization of a unique repeat within a single or small group of chromosomes, compared to the rate of genomic spread, may lead to the formation of specific subsets or subfamilies
(Jorgensen et al., 1986). A similarity in sequence, length of repeat, and centromeric location has been demonstrated for both plant and animal alphoids (Grellet et al., 1986). The persistence of these centromerically linked sequences throughout the eucaryotes, including sequence similarities to the yeast centromeric sequence (Stinchcomb et al., 1985) gives strong indication as to the functional capability of the repeat.

In situ hybridization studies have found a preferential localization of the alphoid family to the centromeric hetero- chromatin regions of metaphase chromosomes (Manuelidis, 1976). This localization has been documented in cells of both animals and higher plants (Grellet et al., 1986; Manuelidis, 1976).

The centromere plays a vital role in the cellular

processes of mitosis and meiosis. The term centromere is applied to a functional and structural region of the chromosome. The functional centromere serves as the point for spindle attachment during metaphase in mitosis as well as metaphase I & II during meiosis. In higher eucaryotes multiple spindle microtubules attach to a visible centromeric trilaminar chromatin structure termed the kinetochore (Jokelainen, 1967; Comings et al., 1971; Roos, 1973; Ris et al., 1981.). Lower eucaryotes (e.g. S. cerevisiae) utilize a single spindle microtubule mechanism, yet do without a structurally distinct kinetochore (Peterson & Ris, 1976). Nuclease digestion experiments in yeast have established a 250 bp resistant core enveloping the centromeric DNA and may represent a primitive kinetochore protein-DNA complex (Bloom & Carbon, 1982).

My work, as presented in this thesis, involved the characterization of the heterogeneity within the human alphoid family. In the process of the characterization, I have described the restriction maps of two 680 bp alphoid clones. I have also determined hybridization conditions which permitted the discrimination between the closely related alphoid families and subfamilies (as defined in this text). These conditions were then implemented in the determination of genomic organization and estimation of copy number. The alphoid sequence homology to the yeast

centromere was investigated as was the ability of the human alphoid sequence to function as a centromere in the yeast S. cerevisiae.

REVIEW OF LITERATURE

The EcoRI Alphoid Family

Digestion of the human genome, with various restriction endonucleases, has revealed a number of tandemly repetitive sequences. Digestion of the human genome with EcoRI reveals two prominent bands representing fragment lengths of 340 and 680 bp (Manuelidis, 1976). About 2% of the total human genome is contained within these two fragments (Wu & Manuelidis, 1980). The two fragments are not present in equal numbers; the 340 bp fragment making up about 2/3 of the total, the 680 bp fragment accounting for 1/3 (Shmookler et al., 1985). *In situ* hybrid- ization studies show a preferential localization of these "alphoid" sequences to the centromeric regions of metaphase chromosomes (Manuelidis, 1978).

To better characterize these prominent tandem repeats, sequencing studies have been carried out. A consensus sequence was determined by Wu and Maneulidis (1980). Uncloned 340 and 680 bp EcoRI generated fragments were sequenced with the resulting sequence representing an average for each base position (Wu & Manuelidis, 1980). The results showed the 340 bp fragment to be comprised of two tandem 169 and 171 bp fundamental repeats (Wu & Manuelidis, 1980). The two repeats do not seem to have evolved from simple (6 to 12 bp) satellite DNA (Wu & Mannuelidis, 1980). The 170 bp monomer is 65%

homologous with the African Green Monkey (AGM) 170 bp monomer (Wu & Manuelidis, 1980). Though alphoid repeat was originally characterized in the AGM, the use of the term alphoid has since been applied to identify the human repeat following the observation of sequence homology. In the dimer each unit shows a 27% variation with respect to each other (Wu & Manuelidis, 1980). The 680 bp fragment contains two tandem 340 bp dimers with the central EcoRI site missing. The two 340 bp dimers making up the 680 bp alphoid have been reported to vary only about 1 % with respect to each other (Wu & Manuelidis, 1980).

Relatedness Among Alphoid Sequences

The first tandemly repeated alphoid sequence was described in the AGM genome (Maio, 1971). Since then, the 172 bp proto- typical AGM repeat has been demonstrated to exist as a related sequence in many eucaryotes. While the human alphoid sequence exists as alternating 169 and 171 bp repeats, found primarily in dimer form (Wu & Manuelidis, 1980), the rat alphoid exists as 185 and 370 bp repeats (Whitney & Furano, 1983). Plants also demonstrate similar repeat lengths; 185 bp maize (Peacock et al., 1981); 352 bp melon (Brennicke & Hemleben, 1983); and a 177 bp repeat from a radish (Grellet et al., 1986) have all been sequenced. Not only are the lengths seemingly conserved, but significant sequence homology has also been detected (Grellet et al., 1986). The similarity

of these characteristics running through both human and non-human eucaryotes has stimulated inquiries into the function of the alphoid families (Wolfe et al., 1985).

Divergence Within the Alphoid Family

Several investigations have been carried out to examine the extent to which the human alphoid consensus sequence represents the repeat sequences occuring throughout the alphoid family. Complete sequencing of several clones has shown a great deal of divergence between EcoRI clones and the consensus. A study carried out by Furlong et al. (1986) compared the consensus sequence to 24 human EcoRI 340 bp alphoid clones. No two of the 24 clones described had an identical sequence when compared to each other nor to the 340 bp consensus (Furlong et al., 1986). The clones had an average sequence divergence of 5.2% from the consensus (an average of 18 base differences per 340 bp), differing in homology from 1 to 41 sites (Furlong et al., 1986). Interestingly, the reported variances occurred in higher frequencies at certain positions than at others indicating that substitutions are occurring in a non random fashion (Furlong et al., 1986). Furlong et al. (1986), reported that the occurrence of a substitution at the same position in more than four of the 24 clones studied was too improbable to justify an assumption of randomness. They present examples of a substitution occurring at a single site (site #157) in

16 of the 24 clones investigated. Another group of 45 EcoRI alphoid clones (all 340 bp and under) were similarly investigated by Jorgensen et al.(1986). They found the degree of divergence between repeats to be greater than first reported within the consensus, ranging from 0.6 to 24.6% in the first monomer and from 0 to 17.8% for the second monomer in the repeat. Three tetramers of the 170 bp repeat (680 bp) have also been tested for consensus sequence homology (Shmookler et al., 1985). Again the divergence was found to be much greater than the consensus sequence revealed. Adjacent 170 bp subunits varied in sequence from each other by 30 to 45% (Shmookler et al., 1985), while the 340 bp dimer subunits varied by 13 to 20%.

Sequencing techniques and restriction digest studies have established that the alphoid sequences possess divergence between species, and have elucidated possible mechanisms responsible for the species-specific self-homogenization from a simple parental unit (Southern 1975a,; Biro et al., 1975). A two step scheme is held as the minimum for the explanation of tandem repeat generation (Wu & Manuelidis, 1980) (Figure 1). The fundamental unit of repetition is duplicated to yield a dimer either through a redundant replication or aberrant crossing-over episode. With time the dimer may acquire sequence divergence with respect to the original monomeric sequence through point mutation. The dimer may then be duplicated and rejoined through amplification and saltation

(Wu & Manuelidis, 1980).

Chromosomal Specific Alphoids

In many mammals, uncloned alphoid DNA has been detected to hybridize to the centromeric regions of all chromosomes (Wolfe et al., 1985). In the mouse , AGM, and Bonnet Monkey, cloned satellite DNA hybridized to the centromeric regions of all chromosomes with no distinction made between chromosomes (Pardue et al., 1970; Kurnit & Maio, 1973; Rubin et al., 1980; Wolfe et al., 1985). While the human alphoid 340 bp sequence is found at all human centromeres (Manuelidis, 1978), the centromeric organization of other alphoid repeat sequences in humans has been demonstrated to be chromosome specific (Mitchell et al., 1985). Certain human chromosomes can now be characterized by their centromeric specific alphoid subfamily sequence and under appropriate experimental conditions (e.g. hybridization stringencies), cloned members of alphoid families can serve as chromosome specific DNA markers (Willard, 1985).

Two EcoRI alphoid fragments of 5.7 and 6.0 kb have been found which are characteristic of the Y chromosome (Smith & Brown, 1987). Low stringency conditions (e.g. low temperature, high salt) allow cross-hybridization of these fragments to autosomes 13,14, and 15, but these chromosomes lacked detectable copies of the 5.7 and 6.0 kb EcoRI repeat sequences (Wolfe et al.,1985; Smith & Brown, 1987). These

repeats, consisting of 170 bp subunits, are found in one major block on chromosome Y (Smith & Brown, 1987).

The X chromosome is characterized by a 2 kb BamHI generated alphoid repeat (Willard et al., 1985). The 12 tandem yet divergent monomers are 60-75% identical in sequence to the consensus (Willard et al., 1985). Sequence comparisons of the 12 monomers has led to the development of a consensus sequence specific for the X chromosome (Willard et al., 1985). The X chromosome has been estimated to carry between 20 to 30 times the amount of alphoid DNA carried on the Y chromosome (Smith & Brown, 1987).

Using the X chromosome's 2kb BamHI alphoid clone as a probe, under variant stringency conditions, researchers have determined alphoid organization on human autosomes. Chromosome 3 contains HindIII generated fragments of 2.7 and 3.0 kb; chromosome 4 contains EcoRI fragments of 1.4 and >25 kb (Willard, 1985). When a similar experiment was carried out at an increased hybridization stringency (allowing 1 to 3% sequence mismatch) most of the non-X-linked fragments were no longer detectable (Willard, 1985) indicating the degree of alphoid sequence divergence between the X chromosome and autosomes 3 and 4.

Other alphoid sequences appear to be less chromosome specific, and are perhaps more typical of mammalian alphoid DNA (Mitchell et al., 1985). Mitchell et al. have described a fragment generated by an EcoRI, HindIII double digest that

is comprised of 14 tandemly repeated variants of a basic 172 bp sequence (termed p82H) (Mitchell et al., 1985). The monomers of this repeat vary an average of 44% with respect to subunit I of the consensus, and an average of 52% with respect to subunit II (Mitchell et al., 1985). The p82H has been found to hybridize to the centromeres of all autosomes and to the centromeres of the X and Y chromosomes (Mitchell et al., 1985). Chromosome 14 showed the highest complement, while chromosome 8 the lowest (Mitchell et al., 1985). Each chromosome possessed distinguishable hybridization patterns, yet bands of 340 and 680 were detected in all chromosomes (Mitchell et al., 1985).

Polymorphisms in Alphoid Sequences

Another example of an alphoid which hybridizes to the centromeres of all human chromosomes is the 3 kb BamHI generated 308 fragment (Jabs et al., 1986). The sequence appears to be enriched on chromosome 6 and has chromosome specific organization on chromosomes 3, 6, 7, 14, X, and Y (Jabs et al., 1986). Organizational polymorphisms of specific chromosomes among normal individuals are also reported (Jabs et al., 1986). In some instances fragments homologous to the 308 appear abundant in one individual and undetectable in another (Jabs et al., 1986). Though the organization of this alphoid may vary between individuals, the total amount in the genome remains relatively constant (Jabs et al., 1986).

Alphoid Function

No translational product is known to be generated from the alphoid families (Potter & Jones, 1983). Using cloned probes, human alphoid DNA was shown to be highly methylated and transcriptionaly inactive (Shmookler et al., 1985). Despite the extent of the documented sequence divergence among the alphoids, Shmookler et al. (1985) have reported regions in cloned 170 bp human repeats, that contain significantly fewer total base changes than would be expected from a calculated average. These regions conform in location to two of three similar sites in the AGM that have a high affinity for a non-histone nuclear protein (Shmookler et al., 1985). This binding protein may play a critical role in nucleosome positioning (Strauss & Varshavsky, 1984). Other proposals regarding alphoid function include their role in chromosome pairing and segregation, speciation and recombination (McCutchan et al., 1982). Wu and Manuelidis (1980) believe the functional contribution by the alphoids to be related to the nature of overall length and tandem repetition rather than individual repeat length, DNA sequence, or secondary structure potential.

In the yeast S. cerevisiae, it has been possible to isolate and clone the chromosomal component that comprises the functional centromere (Stinchcomb et al., 1982). DNA segments were tested for function by their ability to confer mitotic

stability to previously unstable plasmids. The functional sequence (termed CEN), contains a sequence structure that is found once on each of the S. cerevisiae's 17 chromosomes (Fitzgerald-Hayes et al., 1982; Hieter et al., 1985). The recently redefined S.cere-visiae centromere DNA sequence elements (CDEs) are as follows 5'-> 3': CDEI, an 8 bp domain; followed by CDEII a 78 to 86 bp region that is >90% A and T rich; this is followed by CDEIII, a 25 bp domain (Heiter et al., 1985). Ten of the 17 S. cerevisiae chromosomes have had their CEN's sequenced. Of the ten clones described none were identical in sequence (Heiter et al., 1985). The CEN's CDEI and CDEIII are highly conserved (Heiter et al., 1985). The A and T rich region may allow for the formation of a helical configuration, which would distinguish this region from surrounding sequences and influence protein binding or chromatin structure (Fitzgerald-Hayes et al., 1982). Yeast chromatin is organized into nucleosomes in a manner similar to higher eucaryotes (Lohr & Van Holde, 1975). The tripartite organization of the CEN conforms to the postulated structure of regions recognized by DNA binding proteins; a 10 to 20 bp recognition sequence distributed in the DNA at approximate 80bp intervals (Wientraub, 1980).

Significant homologies are also observed between the yeast sequences flanking the high A and T regions and various satellite DNA (including the alphoid) from higher eucaryotes. For example, within the 23 bp Sau Bovine satellite is an 8 bp

sequence exactly homologous to the CDEI on CEN1; another example is an 86% (7 out of 8) homology between a region on the Drosophila melanogaster 359 bp satellite and the CDEI on CEN1 (personal observation). Fitzgerald-Hayes et al. (1982) report a 72% homology between a stretch of human alphoid DNA and element I from CEN 3. No repeated DNA sequences occur in the S. cerevisiae centromere although 30 copies of a 6 kb repeat (Ty elements) have been reported to exist elsewhere in the genome (Clarke et al., 1986). Fitzgerald-Hayes et al. (1982) have investigated the CEN sequence in all three reading frames. The longest open reading frame allows for the coding of a 52 amino acid chain in CEN3 (on yeast chromosome 3) and 40 amino acids in CEN11 (on yeast chromosome 11) (Fitzgerald-Hayes, 1982). They have also determined CEN3 and CEN11 to lack the sequence involved in the initiation and termination of RNA polymerase II transcription (Fitzgerald-Hayes, 1982).

If this single CEN structure per chromosome is responsible for acting as the site for the single spindle attachment, it may be extrapolated that higher eucaryotes require multiple centromeric DNA sequences to accommodate the multiple spindle fibers (Hieter et al., 1985).

The fission yeast Schizosaccharomyces pombe shares a more obvious relationship with higher eucaryotes. It contains 3 chromosomes all larger than the S. cerevisiae, which are visible under light microscopy (Robinson, 1977; Umesono et al., 1983). The centromere region has been detected in two of

the chromo- somes, revealed through chromosome walking (Clarke et al., 1986). The major repeat is 6.4 kb with the total repeat length about 30 kb (Clarke et al., 1986). Within the 6.4 kb repeat are two classes of small repeated sequences of 6 and 12 bp (Clarke et al., 1986). This type of repeat is similar to the simple satellite DNA found in higher eucaryotes.

MATERIALS AND METHODS

Alphoid DNA Clones

Previous work in our lab involved isolating members from several alphoid families and cloning them in pBR322. These include the EcoRI generated 680 bp (pHE340-76, pHE340-77) frag- ments (Doering et al., 1986), the HindIII generated 550 bp (pHH550-31) fragment (Doering & Burket, 1985), and EcoRI generated 340 bp (pHE340-9, pHE340-30, pHE340-64) fragments. The cloned BamHI 2kb (pXBR-1) fragment was a gift from Yang et al., (1982).

Restriction Mapping

All enzymes used in the map construction were purchased from Bethesda Research Laboratories; the digests were performed in the manufacturer's recommended buffers. Unless otherwise noted, the reactions contained 2 ug of plasmid DNA, digested with 5 units per ug DNA of enzyme. Incubation at 37°C for 2.5 hours yielded complete digestion. Reactions were stopped by the addition of SDS and EDTA to final concentration of 0.5% and 10mM, respectively.

Separation of DNA fragments was carried out (depending on fragment sizes) on either 1% agarose or 4% polyacrylamide vertical gels.

DNA Transfers

The DNA fragments were size fractionated by

electrophoresis on 1% agarose gels. The DNA was transferred by the alkaline method (Reed And Mann, 1985) onto Gene Screen Plus membranes (NEN).

Probes

All probes were ^{32}P-labelled by nick translation (Rigby et al., 1977).

Stringency Conditions

All Gene Screen Plus membranes were pre-hybridized at 37°C for 5 hours in hybridization solution (50% formamide, 1 M NaCl, 50 mM Tris (pH 7.5), 1% SDS, 10ug/ml denatured E.coli DNA).

The hybridization temperature varied depending on the percentage of mismatch allowed between probe and membrane bound alphoid DNA. The higher the hybridization temperature the greater the amount of homology necessary to establish a hybrid. The temperatures ranged from 37°C to 55°C. Length of hybridi- zation ranged from 16 to 28 hours.

All membranes were processed with three washes each performed twice. The first consisted of a 10 minute wash at room temperature in 2x SSC (1x SSC is 150mM NaCl, 15mM sodium citrate, 0.1mM EDTA). The second wash consisted of a 30 minute 65°C wash. SDS concentration was held at 1% for all stringencies. Cross-hybridization between alphoids with a greater degree of homology could be reduced with a reduction in SSC concentration. SSC concentration ranged from 2x to

0.001x. The third washes were carried out for 30 minutes at room temperature in 0.1x SSC regardless of stringency.

Membranes were air dried and exposed to Kodak XAR film. Depending on amount of radioactivity remaining on the membranes after the washes, Cronex (DuPont) intensifying screens may have been used to increase the signal.

Copy Number

Each alphoid clone was linearized using an enzyme which has a single site outside of the insert. For the pXBR-1, HindIII was used. All other alphoid plasmids were digested using BamHI. The amount of linear plasmid transferred to Gene Screen Plus was calculated to contain the same number of inserts as would be present in 5 ug of genomic DNA if the individual alphoid sequence occurred in 100, 500, 1000, 2000, or 5000 copies as previously described by Rosenthal and Doering (1983). Genomic DNA was heat treated to accommodate transfer (65°C, 30 minutes). Transfer of plasmid and genomic DNA to the Gene Screen Plus was carried out in 1 ml of transfer solution (0.4 M NaOH, 0.6 M NaCl) using a Schleicher and Schuell Minifold dot blot apparatus. The membrane was hybridized to a radioactively labeled insert-only-probe (insert isolation procedure performed by Eleni Palamidis). This eliminated cross-hybridization to the bound pBR322 from the linearized alphoid clones. Stringency conditions were followed to ensure that alphoid subfamilies did not cross-hybridize. Copy number analysis of each alphoid family and

subfamily was obtained both visually by the comparison of related intensities and by the construction of a standard curve of radioactivity bound to the filter as assayed by a scintillation counter.

Insert Ligation and Amplification

To determine the functional capability of the alphoid sequence we ligated a 340 bp (pHE340-64) alphoid sequence into the yeast shuttle vector pRM47 (Figure 2) (gift from Robert E. Malone). The ligation reaction contained 0.2 ug of EcoRI pRM47, 0.3 ug EcoRI digested pHE340-64, 40 units T4 DNA ligase and ligation buffer (1x is 50 mM Tris (pH 7.5), 10 mM $MgCl_2$, 20 mM DTT, 1mM ATP) in a final reaction volume of 15 ul. The ligation reaction was incubated at 15°C overnight.

Samples of the ligation mix are used to transform E. coli HB101 cells. E. coli cells were grown in LB medium (1.25 g NaCl/L, 1.25 g Bacto-Yeast extract/L, 2.5 g Bacto-tryptone/L, 10 mM Tris (pH 7.5), 1mM $MgSO_4$, pH adjusted to 7.5, plus ampicillin (50 ug/ml) added at 37°C) to an A_{560}=0.6. Cells were made competent by the following procedure: 10 mls of culture were spun for 5 minutes at 7k in sterile conditions. The pellet was resuspended in 5 mls cold sterile 0.01 M NaCl. Cells were respun and resuspended in 5 mls cold sterile 0.03 M $CaCl_2$. Respun cells were then resuspended in 1 ml cold transformation buffer (70 mM $MnCl_2$, 50 mM $CaCl_2$, 40 mM NaAc). DNA was added in 10 to 15 ul aliquots to 0.2 mls of competent E. coli cells. Heat shock was achieved at 42°C

for 2 minutes. Eight-tenths mls of cold L-broth was added to each sample. Samples were then plated onto L-plates containing ampicillin and a gridded nitrocellulose filter. A colony lysis was performed on ampicillin resistant cells by allowing the gridded filter to absorb 0.5 M NaOH for 7 minutes, 1 M Tris (pH 7.5) for 2 minutes (twice), and 2x SET for 4 minutes. The filter was air dried and quick-dipped in 95% EtOH. The filter was dried at 60°C overnight. pRM47-340 transformants were detected by colony hybridization (Grunstein & Hogness, 1975) with total human genomic DNA as probe. A single positive transformant (pRM47340-1) was selected (Figure 3).

Yeast Transformation

pRM47 and pRM47340-1 DNA was then used to transform S. cerevisae strain RM13-4B, (also a gift from R. Malone). The yeast were grown to a concentration of $1\text{-}3x(10)^7$ in 40 mls of YPD (10 g Dextrose, 10 g Bacto-peptone, 5 g Bacto-yeast extract, in 500 mls distilled H_20). Cells were spun at room temperature and washed in 10 mls TE (10 mM Tris pH 8, 1 mM EDTA, pH to 7.5). After repeating the TE wash, cells were resuspended in TE to a concentration of $2x(10)^8$ cells/ml. Cell numbers were determined by a serial dilution performed in phosphate buffer (14.2 g $NaHPO_4$ up to 500 mls distilled H_2O, pH to 7.5) and plating on complete media (20 g Bacto-dextrose, 1.7 g Yeast nitrogen base, 5 g NH_4SO_4, 30 to 35 g

Bacto-agar for use in plates, brought up to a volume of 900 mls with H_2O, pH adjusted to 5.8 and autoclaved. After cooling (65° C) 10 mg adenine-sulfate, 20 mg uracil, 50 mg arginine-HCl, 20 mg histidine-HCl, 80 mg homoserine, 100 mg leucine, 50 mg lysine-HCl, 20 mg methionine, 50 mg phenylalanine, 50 mg tryptophan, 50 mg tyrosine, 75 mg aspartic acid, 50 mg isoleucine, 100 mg threonine, 150 mg valine were added and the suspension brought up to a final volume of 1000 mls with distilled H_2O. Five-tenths mls of cell suspension was transferred to a 12 ml snap top tube and 0.5 ml of 0.1 M LiCl were added. After shaking for 1.5 hours at 30°C (100 rpm) 0.1 ml was transferred to a sterile microfuge tube. The pRM47 or pRM47340-1 DNA was added and allowed to sit for 30 minutes at 30°C. An equal volume of filter sterilized 70 % PEG-4000 (polyethylene-glycol) was added and the mixture vortexed and allowed to stand 1 hour at 30°C. The heat shock was carried out at 42°C for 6 minutes and immediately cooled to 20°C for 3 minutes. Cells were spun down in microfuge tubes for 1.5 minutes at 12K. The pellet was resuspended in 1 ml sterile H_2O. 2, 10, 50, and 250 ul were plated on selective plates (complete media without uracil substituent). One-tenth ml of a 10^{-3}, and 0.1 and 0.3 mls of a 10^{-4} phosphate buffer dilution were plated on complete media, and 0.3 mls of a 10^{-4} dilution were plated on YPD plates (YPD and 8.5 g agarose).

Plasmid Stability Assay

The relative stability of the pRM47 and the pRM47340-1 transformants was compared. One colony from each strain of transformants were placed in 1 ml liquid YPD and vortexed gently before adding to 25 mls of non-selective YPD media (pre-warmed to 30°C) and shaking at 250 r.p.m.. Aliquots were taken immediately and at several hour intervals, and plated on uracil deficient and complete media. To ensure no essential metabolite was carried in the drawn off aliquot, the liquid media was eliminated. Cell suspensions were spun at 5K for 1 minute. The pellet was then resuspended in phosphate buffer and sonicated on an Artek 300 dismembrator for 15 seconds at 30% power. The concentration of cells was determined using a hemacytometer and proper dilutions were subsequently performed to maximize plating efficiency (a goal of 30 to 300 cells per plate). Plated cells were inverted and incubated at 30°C from 36 to 72 hours. The percentage of cells that retained the plasmid was determined as the ratio of the number of cells that had retained the URA3 gene conferred by the pRM47 plasmid (grew on uracil deficient plates) over the total number of cells (growth on YPD). The deletion in the yeast genome of a URA locus ensured that no integration of the URA containing plasmid could occur._

Alphoid and CEN Comparison

The initial investigation employed an Apple compatible

search program that was capable of screening the human 340 bp alphoid consensus for sequence relatedness to elements I and III of the S. cerevisae's CEN sequence. Subsequent visual compari- sons were done to include the CEN`s redefined CDEI and CDEIII. Positions on the alphoid consensus that showed significant homology to the CDE elements were then investigated in non-consensus alphoid variant sequences for an increase in total percent homology.

RESULTS

Restriction Maps

The first experiment was designed to determine the variability in sequence between two 680 bp alphoid clones (pHE340-76,pHE340-77) and the 680 bp consensus (Wu & Maneulidis, 1980). This was investigated by performing single and double digests of the two plasmids and comparing their restriction maps to each other as well as to the consensus sequence restriction map. Using the same restriction enzymes on pHE340-76 and pHE340-77 as well as plasmid alone (pBR322) allowed for the identifi- cation of insert-containing fragments. The lengths of the unique fragments were calibrated using the known fragment sizes of digested pBR322. Restriction maps of the two clones were com- pared to the consensus map, revealing several variations (Figure 6). The pHE340-76 has a missing Sau3A positioned at 52 bp, and at 392 bp. The pHE340-76 has an additional HaeIII site at 125 bp and an additional AluI site at 340 bp. The pHE340-77 has a missing HinfI site at 172 bp, a missing Sau3A site at 392 bp, and a missing HaeIII site at 636 bp. The pHE340-77 has an additional HinfI site positioned at 263 bp and an additional HindIII/AluI site at 340 bp.

Elimination of Alphoid Cross Hybridization

To further test relatedness among alphoids, a stringency condition protocol effective during hybridizations was established as washes, which allowed the elimination of cross-hybridization between seven different alphoid clones. Once these conditions were determined for each alphoid, they could then be used as a tool in investigations concerning organization and copy number in the human genome.

A low stringency condition (hybridization at 37°C, wash at 2x SSC/ 1% SDS 65°C) produced cross-hybridization between both of the EcoRI 680 bp fragments (pHE340-76,pHE340-77). The two 680 bp clones also cross hybridized to the 340 bp pHE340-64, 2 kb pXBR
1, and weakly to the 550 bp pHH520-31 (see Figure 7). An increase in stringency was obtained by lowering the SSC concentration to 0.1x. Figure 8 shows that hybridization with the pHE340-76 at this increased stringency eliminated cross-hybrid-ization to the pXBR-1 and pHH520-31. At this stringency the pHE340-76 still cross-hybridized to the pHE340-77 and pHE340-64.

Two additional 340 bp alphoid clones (pHE340-9, pHE340-30) were added to the study. Elimination of cross-hybridization between the two 680 bp and one of the 340 bp EcoRI alphoid clones occurred when the hybridization temperature was increased to 50°C while keeping the SSC wash

concentration constant at 0.1x. At this stringency a minimum of 95 % homology must be present for hybridization to occur (previous work in our lab by E. Palamidis). At this stringency four different alphoid clones became specific only for sequences at least 95% homologous with themselves (i.e. pHE340-76, pHE340-64, pHH550-31, and the pBRX-1) demonstratable by the elimination of cross hybridization (Figure 9). At this same stringency the pHE340-77, pHE340-30 and the pHE340-9 were found to cross-hybridize to each other (Figure 10). Elimination of the pHE340-9 cross-hybridization to the pHE340-30 and the pHE340-77 was achieved when the stringency was increased by raising the hybridization temperature to 55°C and the wash condition to 0.04x SSC (Figure 11). Cross hybridization between the last two clones (pHE340-30, pHE340-77) persisted until the SSC was lowered 20 fold (hybridization temperature 55°C, wash 0.002x SSC) (Figure 12). An adjusted exposure time would eliminate the residual intensity detected. Confidence in our ability to be able to distinguish subfamilies is based in the degree of stringency necessary to eliminate cross hybridization between these differing members. In addition to stringency requirements, exposure time of the blots to film may also play a role in visualization of any cross hybridization.

Copy Number

Gene Screen Plus containing the transferred genomic and

alphoid DNA was subjected to the specific stringency which selectively allowed the hybridization of specific alphoid family and subfamily probes. Autoradiographs of the hybridization show degrees of intensities which are directly related to the amount of bound radioactivity (Figure 13). When the autoradiographs reveal a low level of background radiation, the regions of hybridization were cut out of the Gene Screen and assayed for bound radioactivity using liquid scintillation counting. If interference from non-specific hybridizations occurred, the regions of DNA transfer could not be accurately assayed. In this case only the visual assessment of intensity comparison was performed.

Figure 13 is the autoradiograph of the copy number assay of a 340 bp alphoid clone (pHE340-64). The intensity of the genomic DNA appears to fall between the region of intensity representing 5000 and 2000 copies. A standard curve produced by the scintillation counting of the alphoid hybrid intensities, estimated the genomic copy of the pHE340-64 to be 1900. A similar assay for the 520 bp alphoid clone (pHH550-31) established the genomic copy number of this sequence to be 8100 (figure not shown). A visual comparison of the autoradiograph intensities for the BamHI 2 kb alphoid (pXBR-1) revealed a genomic intensity that is somewhat higher than the 5000 copy standard (figure not shown).

Genomic Organization

Genomic DNA was digested to completion by restriction endonucleases that do not cut within the 680 bp consensus. These are AluI, HindIII, EcoRI, and BamHI. Blots were constructed with 4 lanes, each lane containing DNA digested with a different enzyme. Determination of the genomic organization for sequences closely related to the clones described, necessitates the detection of each class of alphoid sequence in the genome without risk of cross-hybridization between the sequences that occur in the various alphoid families and subfamilies. Such stringencies were described above.

A comparison of organization between the two 680 bp alphoids is shown in Figure 14. This low stringency hybridization does not eliminate cross-hybridization between the two 680 bp clones, yet differences in band pattern between the two can be seen. Several examples are the intense 2 kb AluI and 3.3 kb BamHI bands sharing a greater homology to the pHE340-77. The pHE340-76 shows homology to some high molecular weight BamHI fragments that are not detected by the pHE340-77. The homology to fragments produced by HindIII is similar except for a 3.3 kb fragment which shares a greater homology to the pHE340-76. Both clones hybridize to the EcoRI generated 340 and 680 bp fragments.

Organizational differences were more clearly defined when

using the specific stringencies for elimination of cross-hybridization between the variant alphoid sequences. The increase in stringency necessitated a 10 fold increase in DNA to elucidate the bands of hybridization. The obvious difference is the reduction of total bands, due to an increase in specificity as seen on Figure 15, which shows the organizational pattern of two 680 bp EcoRI alphoid clones (pHE340-76, pHE340-77). The pHE340-77 clone is specific at this stringency and hybridizes to the same BamHI 3.3 kb fragment detected at the lower non-specific stringency (Figure 14). Both 680 bp clones share homology with the 340 bp EcoRI band but virtually none is detected to the 680 bp EcoRI band. Figure 15 also shows the pHE340-77 sharing homology with two AluI fragments (1.7 and 0.85 kb) not detected by the alternate 680 bp pHE340-76. The pHE340-76 demonstrates some detectable homology to HindIII fragments which are not detected by the pHE340-77. The organization of the 340 bp repeat family (Figure 16) is distinct from that of the 680 bp EcoRI family (Figure 15). The clearest example is the loss of hybridization of the 3.3 and 1.7 kb pHE340-77-like fragments by both 340 bp clones. Both clones hybridize to the 340 bp fragments more intensely than the 680 bp fragments. The pHE340-64 hybridizes to a detectable high molecular weight HindIII repeat while the pHE340-9 does not (Figure 16). Another variant pattern of hybridization is demonstrated for the three distinct families (the pHE340-77, pHH550-31, and the

pXBR-1) by Figure 17. Each of the families hybridize to specific size fragments. The pHH550-31 hybridizes to 1 kb EcoRI and AluI genomic fragments, and a series of lower molecular weight HindIII fragments. The pXBR-1 hybridizes to its own 2 kb BamHI fragment very intensely as well as to unique 4.5 kb BamHI fragments. Concentration of pXBR-1 like sequences are also found in 850 and 680 bp Alu fragments.

Human Alphoid Function in Yeast

Exogenous autonomously replicating plasmids in transformed yeast are lost at a high frequency when grown in a non-selective media. The nucleotide sequence (CEN) that defines the centromere in the yeast S. cerevisiae stabilizes exogenous plasmids in non-selective media (Clarke and Carbon, 1980a, 1980b). A functional relationship to the alphoid was tested in light of the finding of a CEN sequence relationship to the human alphoid consensus (Fitzgerald-Hayes et al., 1982).

A representative of the EcoRI human alphoid family (pHE340-64) was chosen to test the ability of its sequence to confer stability in a non-selective environment. The pHE340-64 was selected for its relative simplicity (340 bp). The alphoid insert was subcloned into the pRM47 EcoRI site (see Figure 2) to form pRM47340. A transformation of HB101 E.coli cells was performed. Ampicillin resistant cells were picked from a master plate and placed on filter paper for a colony

lysis and subsequent screening with nick translated human genomic DNA. Positive colonies were chosen for a plasmid prep (Figure 3). This plasmid as well as the pRM47 alone were used to transform S. cerevisiae strain RM13-4B. RM13-4B is auxotrophic for uracil. The essential uracil contribution from the pRM47 was used as the selectable marker. We were then able to directly compare the mitotic stability of the two new strains while they were being allowed to proliferate in a non-selective broth.

Data from two individual kinetic experiments are presented. Figure 18 represents a pattern of plasmid loss in the two strains. Freshly transformed cells were used as the inoculum. The graph represents 7 points where cells were removed from the liquid media, processed (see Material and Methods), and plated. The retention of plasmid is continuously greater in pRM47 strain, beginning at t=5 hours (63%) and concluding at t=27 hours (6%). The pRM47340 lags at t=5 hours (25%) throughout to t=27 hours (1%). The plasmid is lost rapidly in both strains. The next graph is of a similar experiment (see Figure 19). The inoculum used in this study were cells from a previous stability assay and were not freshly transformed. Both colonies were of a similar size and both were picked off a selective plate. The only obvious difference is that the pRM47 source came off a t=6 hour plating and the pRM47340 source came off a t=12 hour plating; both from the same experiment. The graph shows a peak which

is regarded as an artifact unless the first plating is ignored (t=5). The latter option allows the graph to depict the opposite of the first. The retention is higher at t=8 hours for the pRM47340 at 30% and lower for the pRM47 at 27%. The rapid loss of plasmid is detected for both strains following the t=8 hour plating; the pRM47 again possessing the greater retention from t=12 to t=24 hours.

CEN and Alphoid Comparison

The two functional elements making up the CEN (elements I and III) were compared to the 340 bp alphoid consensus for sequence homology. All three sequences were fed into an Apple search program for analyses. The initial screening was used to examine homology between the 340 bp consensus and the redefined CDE's (Heiter et al., 1985). Figure 4 shows the resulting comparison with the two greatest examples of CEN homology shown superimposed on the 340 bp consensus. These were chosen for their greatest total homology. The first region of homology is the CDE I to the alphoid's consensus at positions 37 to 44 (62.5%). The CDE III is 50% homologous to positions 139 to 163 on the consensus. Another example of homology exists between CDE I and positions 208 to 215 on the consensus as well as CDE III and the alphoids positions 306 to 330 (37.5 and 60.0% respective-ly). As would be expected in a 170 bp repeat, regions on the alphoid representing CDE I were approximately 170 bp apart. A similar distance is seen

(Figure 4) between the regions homologous to the CDE III. A region on the alphoid sequence exists between the paired CDE I and CDE III, of approximately 91 bp that is 60% A and T rich.

The study was extended to include an investigation of the homology between the CDE and sequenced clones of human variant 340's (Figure 5). Jorgensen et al. (1986) described about 30 different sequences. No variant 340 contained a base substitution that would increase the percent homology to the CDE III relative to the 340 bp consensus (considered as a beneficial change). Any clone that contained a single beneficial base change (seen at positions 139 to 163, 306 to 330) also included 3 or 4 non-beneficial substitutions. Several clones were examples of beneficial base substitutions in the regions homologous to the CDE I (Figure 18). The variants examined showed an increased homology to the CDE I, changing it from 62.5% to 75% (5 clones at positions 37 to 44, and 6 clones at positions 208 to 215). A comparison was performed of the CDE to sequences in the 2 kb BamHI and its variants at regions which demonstrated homology to the 340 bp consensus. Figure 5 lists the comparison of the CDE I to the 2 kb BamHI consensus (170 bp), as well as its variants. Six 170 bp monomers from the BamHI 2 kb fragment are presented that possess a greater homology (37.5%) to the CDE I than does the 170 bp consensus for the 2 kb fragment (12.5%). None of the 170 bp repeats making up the consensus for the 2 kb

alphoid sequence increased the homology to the CDE III. Any sequence with a beneficial substitution that would increase the homology also had base changes at other positions that would either decrease the percent homology or keep it the same.

DISCUSSION

Restriction Maps

Much work in the past has dealt with the heterogeneity among 340 bp alphoid clones (Jorgensen, et al., 1986, Furlong, et al., 1986). Work done previously in our lab on total genomic sequences (Doering, et al., 1986) has demonstrated a high degree of variance to also occur in the 680 bp alphoid repeat. I have characterized, by restriction mapping, two 680 bp alphoid clones (Figure 6) and found them to be structurally distinct from each other as well as from the consensus (Wu and Maneulidis, 1980). Furthermore, it appears that the distribution of sequence mismatch is between the two investigated clones and the consensus is random. The finding of variants within the first two clones examined indicates that it is likely the 680 bp is more divergent than reported by Wu and Maneulidis (1980).

Stringency

To examine the genomic organization and copy number of each alphoid family and subfamily represented by available alphoid clones, I developed stringency conditions which enabled me to eliminate any cross-hybridization outside of homologous subfamilies (Figures 7-12). The alphoids have also

been characterized according to the stringency conditions necessary to eliminate the cross-hybridization, permitting me to categorize the relationship between them. Those alphoids which did not cross-hybridize at low stringency (hybridization at 50 C, wash 0.1x SSC at 65 C) were divergent enough to be classified as distinct families (pHE340-76, pHE340-64, pHH550-31, and pXBR-1). The 340 bp pHE340-9 is more related to the pHE340-77 than another 340 (pHE340-64). The remaining EcoRI generated alphoids which required a higher stringency (hybridization at 55 C, wash 0.002x SSC at 65 C) to eliminate cross hybridization were characterized as belonging to the different subfamilies within the same family (pHE340-77, pHE340-30). It is interesting that a 680 bp fragment (pHE340-77) would share greater homology with two 340 bp fragments (pHE340-30, and pHE340-9) than with another 680 bp fragment. The divergence also appears to be great between the pHE340-64 and the other two examined 340 bp fragments (pHE340-30, and pHE340-9) as determined by the relatively low stringency needed to eliminate cross-hybridization. The reasons for the divergences within the 680 bp or 340 bp alphoids may be due to the saltation or amplification of the basic ancestral repeat unit. The divergences may be the result of homogenization of the original unit following a mutational event which was isolated to a single or to a small number of isolated chromosomes, thereby any combination of mutatation amplification would create repeats unique to the involved

chromosomes. Using the stringencies described here, it may be possible to determine the chromosomal organization of these seven alphoid clones. The variances observed between alphoid families and subfamilies may be explained by evolution of these due to their location on different chromosomes.

Copy Number

Through the use of specific stringencies I have been able to screen the human genome for a select group of sequences (Figure 7). The results show that the alphoids do not exist in the genome in similar numbers, but do account for a substantial part of it. The pHH550-31 has been calculated to account for greater than 0.34% of the entire genome (8100 copies x 2040 bp / 3 x 10^9 genomic bp). The pHE340-64 (1900 copies) and pXBR-1 (>5000 copies) have been similarly calculated to account for 0.02% and 0.14% of the genome respectively. As would be expected, the proportion of copies is higher for the alphoids which belong to an entire family (pHH550-31 or pXBR-1) than for the alphoid belonging to a subfamily of the large EcoRI family (pHE340-64). The results for the pXBR-1 are similar to the 6000 copies reported by Yang et al. (1982).

Genomic Organization

Since the similarities between alphoid sequences predispose them to cross-hybridize, I followed the

hybridization procedures as described in the Results section. The first example of genomic organization (Figure 14) shows differences in both the pHE340-76 and pHE340-77 despite both being EcoRI 680 bp fragments. The pHE340-76 and pHE340-77 both hybridize to HindIII generated repeats at this lower stringency condition. Figure 14 shows the organizational pattern of these two clones at the stringency where cross-hybridization is eliminated. The pHE340-76, but not the pHE340-77 exhibits that it has sequence relationships to high molecular weight HindIII fragments. Of greater interest is the drastic reduction of hybridization of both these 680 bp alphoid clones to the 680 bp EcoRI generated genomic fragments at high stringency. The low level of hybridization between the two clones and the genomic 680 bp sequences shows the relatively small proportion of these specific cloned sequences present in the genome. Both show homology to the 340 bp genomic fragments. Similar to the pHE340-76, the pHE340-64 shares homology to high molecular weight HindIII fragments (Figure 16) which the pHE340-9 does not, demonstrating that the HindIII generated fragments are a distinct sequence class onto themselves. The pHE340-9 and pHE340-64 are both highly selective for genomic 340 bp fragments, while the 680 bp clones do not share this selectivity for 680 bp fragments. Thus the 680 bp fragments in the genome may be in fact more divergent than the 340 bp fragments. The differences in genomic organization are best observed between the pHH550-31

and pXBR-1 but structural organizational distinctions between all three of these families is also appreciated (Figure 17). The organization of the pHH550-31 seems to be concentrated in lower molecular fragments; unlike the pXBR-1 where the sequence is distributed in a variety of repeat lengths, with notable concentration at 2 and 0.68 kb. The high copy number of pXBR-1 in the genome is also detected from this blot by the relatively high degree of hybridization intensity.

Alphoid and Cen Comparison

To gain insight as to the conservation of the centromerically linked alphoid sequences, I compared the sequence relationship between the CEN of the yeast S. cerevisiae to alphoids from the human genome (Figures 4 & 5). I found regions on the alphoid sequences that shared both sequence and structural homology to the CEN. After locating the area of highest homology to the CEN (in this case the CDE I) I looked 80 to 90 bp downstream for a region which would represent the second element on the functional CEN, the CDE III. I found two sets of CDE like regions on the alphoids, each set 170 bp apart indicating the size of the repeat sequence. The homology was as high as 75% for the CDE I and 50% for the CDE III with a 90 bp intervening sequence being 60% A and T rich. This very closely resembles the structure of the CEN's 78 to 86 bp intervening sequence though the A and T composition on the CEN reaches >90%.

The sequence comparison included variants of the 340 bp as well as the 2 kb BamHI consensus. I found the 340 bp consensus to share a closer relation to the CEN than the 170 bp repeats within the BamHI 2kb repeat which may mean that the 340 bp human repeat shares a more primitive parent sequence with the yeast than with other divergent human alphoid sequences. This evidence of an evolutionary tie may be used in future assays for centromeric function. The analyses also yielded the finding of a preferential conservation of the CDE I over the CDE III in the alphoid sequences, as demonstrated from the sequence homology comparisons (Figure 5). The variant sequences examined also revealed a conservation of the CDE I, including a substantial number of beneficial base pair substitutions. A CDE I-like sequence may be instrumental as the site for binding proteins or in the formation of secondary structure. It may be wise to give special considerations to the CDE I conservation in future functional and evolutionary investigations.

Function

One of the most intriguing explanations for the high centromeric inclusion of eucaryotic repetitive DNA deals with the possibility of its centromeric function. The functional centromeric component in the budding yeast S. cerevisiae has been isolated, and a relationship to human alphoid DNA has been described (Fitzgerald-Hayes et al., 1982), yet in the

fission yeast S. pombe the CEN remains unresolved. The centromeric DNA in S. pombe is much more complex than S. cerevisiae including 30 kb of heterochromatin-like sequences making it similar to higher eucaryotes (Clarke et al., 1986). The jump in complexity to the centromeres in humans can be expected to be many times larger than that observed between these two species of yeast. Even with this given range in species complexity, evidence exists for an evolutionary bond between yeast and human DNA. With evidence of the CEN and alphoid sequences similarites in hand, I attempted to test a simple representative of the human alphoid sequence for centromeric function. The results presented were inconclusive. The only statement that is intended here is that both plasmids were lost at comparable rates. other than (Figures 18 & 19). No consistent evidence for plasmid stabilization was established. The reasons may lie in the heterogeneity of the system itself. The human alphoid sequence may not be recognized by the yeast system. The S. cerevisiae CEN does not function even in the S. pombe (Clarke L, et al., 1986). If the alphoid sequences are in fact the entity conferring centromeric function upon the cell, it is very likely that the tandemness is important to function and a single sequence might be worthless in demonstrating a role.

Another possibility may have been in the wrong selection of a particular alphoid sequence. Many variant sequences have been determined, and it may be a single sequence within the

mass of variants that confers stability and that base substitutions are not important to function as long as the secondary structure and overall organization is not disrupted. The observation of instability may be due to the size of the plasmid. Perhaps the size of pRM47 alone, with no insert is the limit tolerated by the yeast, and the insertion of the 340 bp alphoid is too large for the organism to retain. An assay to detect instability due to size could use a 340 bp non-alphoid subcloned into the pRM47.

Future experiments could include different alphoid clones, beginning with single repeats and progressing to long tandem stretches. To avoid the heterogenous system described researchers may have to investigate transfected human cell lines.

Figure 1.

Scheme representing the generation of a dimer from simple parental sequence as well as a representation of an interspersed and tandem repeat

Figure 1 is the scheme as described by Manuelidis & Wu (1980). Ancestral sequence is reduntantly duplicated to yield a dimer. Over time this sequence aquires sequence divergence. Unequal crossing-over or other recombinational event may join two such repeats. Amplification results in long stretches of the particular repeat.

45

Fig 1

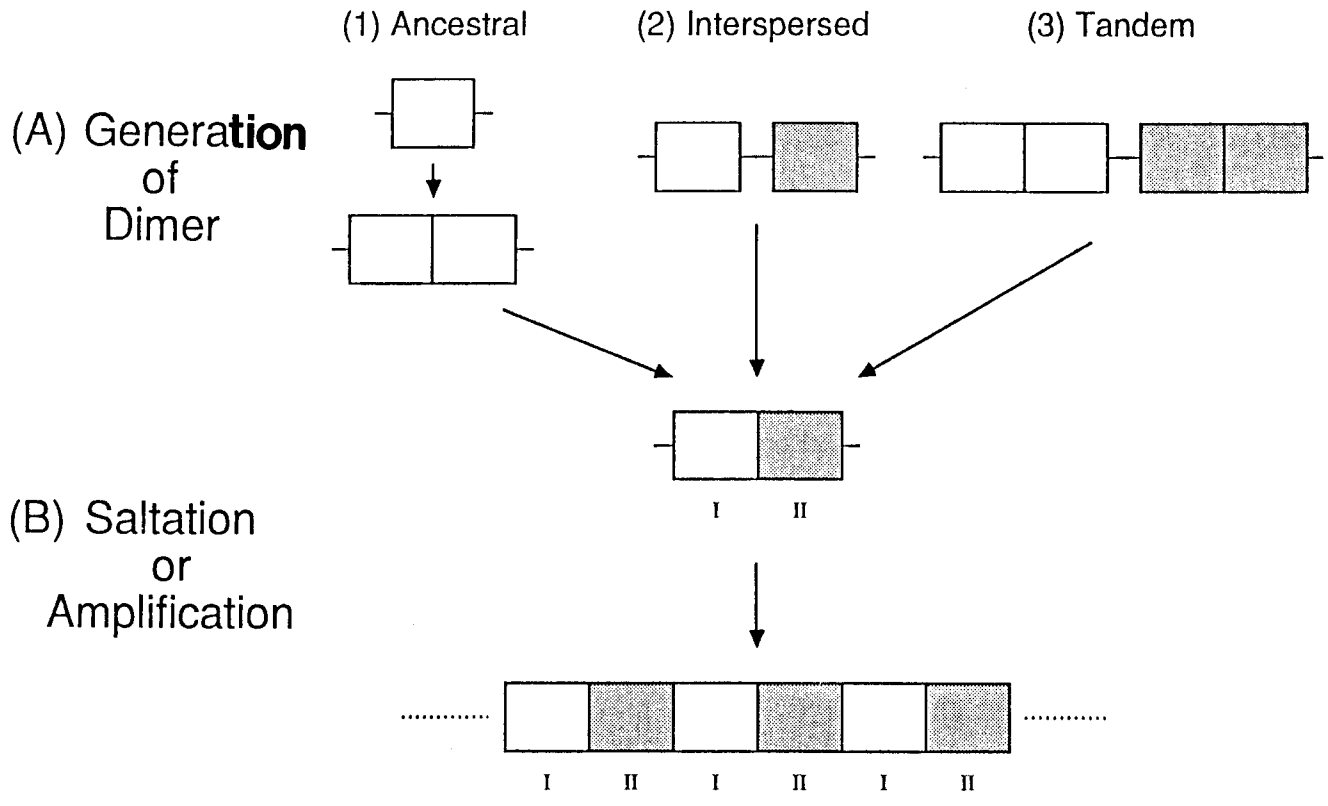

(1) Ancestral
(2) Interspersed
(3) Tandem
(A) Generation of Dimer
(B) Saltation or Amplification
I
II
I
II
I
II
I
II

Figure 2.

Representation of yeast shuttle vector pRM47

The plasmid has the ability to replicate in both bacteria and yeast. It possesses selectable markers for transforming bacteria and yeast. The plasmid has an EcoRI site allowing the ligation of the EcoRI alphoid 340 bp fragment.

ars= autonomously replicating sequence (yeast)
ori= origin of replication (bacteria)
amp= ampicillin resistance gene (bacteria)
URA 3= uracil gene (yeast)
EcoRI= site of ligation

Fig 2

pRM47

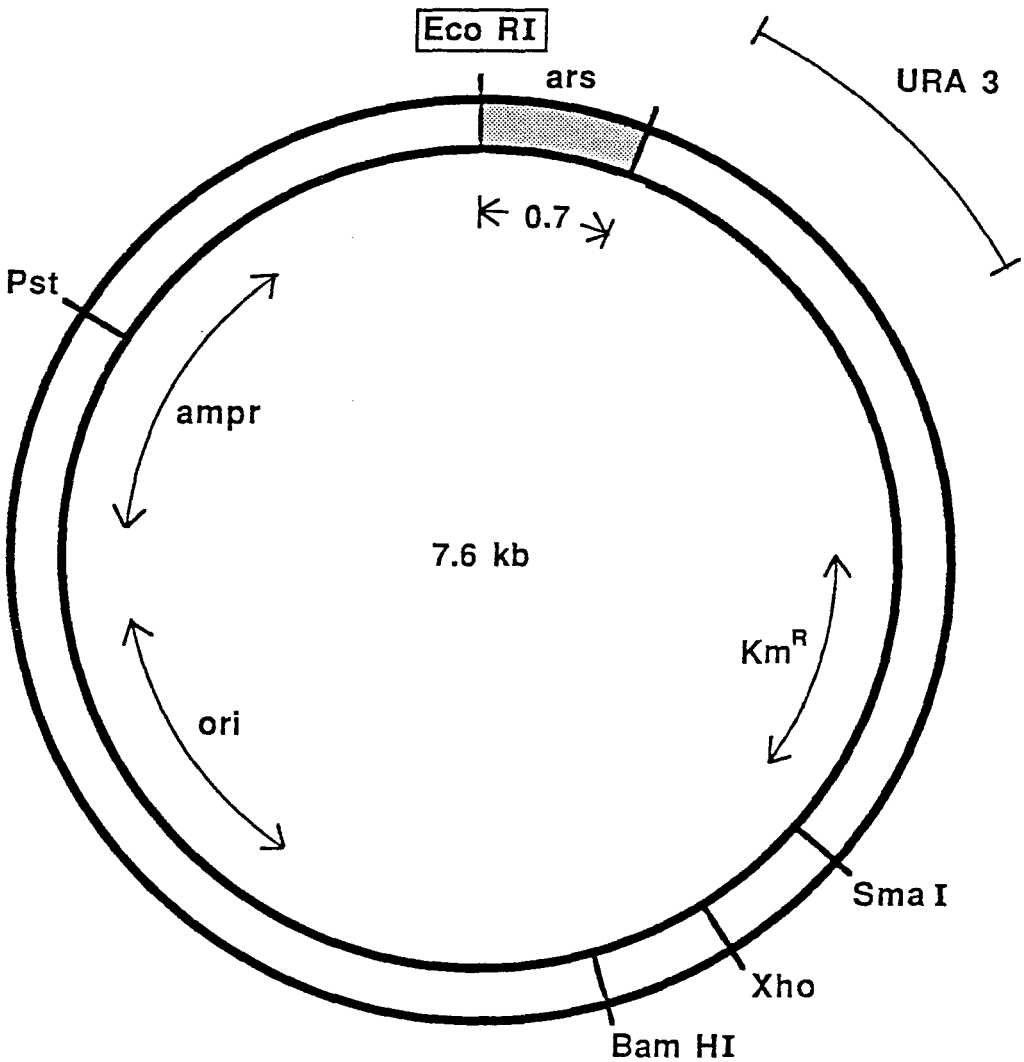


Figure 3.

Colony hybridization of transformed HB101 cells with a pRM47 plasmid with a ligated human 340 bp alphoid clone (pHE340-64)

Ampicillin resistant cells were picked off a master ampicillin containing plate and place on filter. Cells were lysed (Grunstein & Hogness, 1975) and hybridized to total human genomic DNA. Colony #1 was retrieved from master plate for DNA amplification.

Fig 3

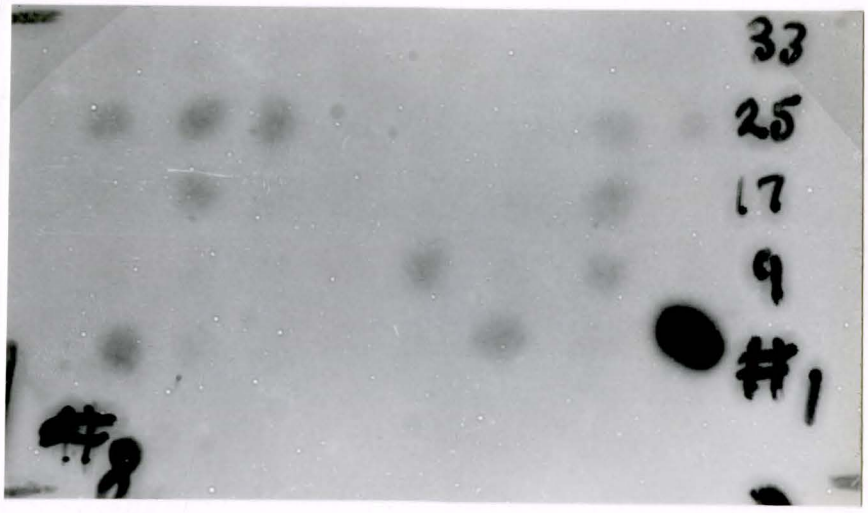


Figure 4.

Comparison of yeast CEN to 340 bp human alphoid consensus sequences

Entire 340 bp consensus sequence is shown with regions of homology to the CDE I & III in bold.

Fig 4

```
        10         20         30         40         50         60
AATTCTCAGTAACTTCCTTGTGTTGTGTGTATTCAACTCACAGAGTTGAACGATCCTTTA
                                      [ CDE I]
        70         80         90        100        110        120
CACAGAGCAGACTTGAAACACTCTTTTTGTGGAATTTGCAAGTGGAGATTTCAGCCGCTT

       130        140        150        160        170        180
TGAGGTCAATGGTAGAATAGGAAATATCTTCCTATAGAAACTAGACAGAATGATTCTCAG
                  [        CDE  III          ]
       190        200        210        220        230        240
AAACTCCTTTGTGATGTGTGCGTTCAACTCACAGAGTTTAACCTTTCTTTTCATAGAGCA
                             [ CDE I]
       250        260        270        280        290        300
GTTAGGAAACACTCTGTTTGTAAAGTCTGCAAGTGGATATTCAGACCTCTTTGAGGCCTT

                  310        320        330        340
            CGTTGGAAACGGGATTTCTTCATATTATGCTAGACAGAAG
                 [         CDE III          ]
```

Figure 5.

Sequence comparison of the CDE I & CDE III to the regions of homology on the consensus of the human alphoid 340 bp and to the same regions of sequenced variants.
Sequence comparison of the CDE I & CDE III to the regions of homology on the consensus of the pXBR-1 2kb BamHI alphoid and to the same regions on sequenced variants .

Fig 5

```
                                          A/   A/
                    CDE I consensus GTCACGTG

positions 37 to 44 in 340 bp
          EcoRI alphoid consensus  C----AGA
                              aR1-131 C-----G-
                              aR1-123 ------GA
                              aR1-13  ------GA
                              aR1-110 ------GA
                              aR1-121 ------GA

positions 37 to 44 in 170 bp repeat of
      2 kb BamHI alphoid consensus TCGTT-GA
                                   aX2 --GTT-GA
                                   aX3 --GGT-GT
                                   aX4 T-GTT-GA
                                   aX7 --GGT-GT
                                   aX8 -CGGTGGT
                                  aX12 --GGT-GA

positions 208 to 215 in 340 bp
           EcoR1 alphoid consensus C-----GA
                              aR1-18 C------A
                              aR1-28 C-----GA
                             aR1-132 T-----G-
                              aR1-32 ------GA
                              aR1-78 ------GA
                             aR1-103 ------GA



                                        A/
                  CDE III consensus TGTTTTTG.TTTCCGAAA....AAA

positions 139 to 163 on 340 bp
                  alphoid consensus A-GAA--A.C----T-T-....CT-

positions 306 to 330 on 340 bp
                  alphoid consensus GAAACGG-.----TTC-T....TGC
```

Figure 6.

Restriction maps of pHE340-76, consensus, and the pHE340-77

The restriction site are are marked and numbered according to the positions of cleavage by the enzymes. Positions are 5'->3' from position 0 (EcoRI). Maps were constructed by calibrating fragments of the digested pHE340-76 and pHE340-77 against known fragment sizes of digested pBR322.

Fig6

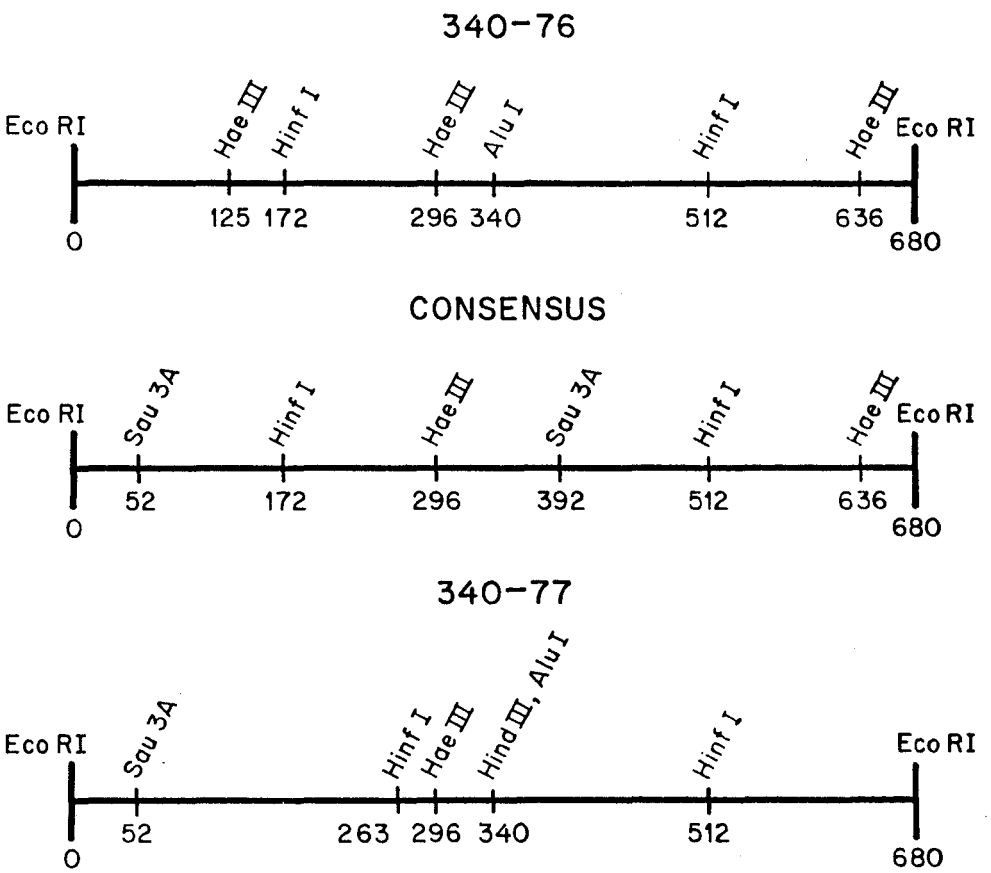

340-76
Eco RI
Hae III
Hinf I
Hae III
Alu I
Hinf I
Hae III
Eco RI
0
125 172
296 340
512
636
680
CONSENSUS
Eco RI
Sau 3A
Hinf I
Hae III
Sau 3A
Hinf I
Hae III
Eco RI
0
52
172
296
392
512
636
680
340-77
Eco RI
Sau 3A
Hinf I
Hae III
Hind III, Alu I
Hinf I
Eco RI
0
52
263 296 340
512
680

Figure 7.

Determination of stringency strengths in the elimination of cross-hybridizations

Five alphoid clones (0.1 ug) were digested to cleave the insert at site of ligation: EcoRI BamHI HindIII

pHE340-64 pXBR-1 PHH550-31
pHE340-76
pHE340-77

Fragments were resolved on a 1% agarose gel and transfered to Gene Screen Plus membranes. The membranes were hybridized as indicated. Hybridization at 37 C, wash at 2x SSC/1% SDS 65 C.

1=pHE340-76
2=pHE340-77
3=pHE340-64
4=pXBR-1
5=pHH550-31

Fig7

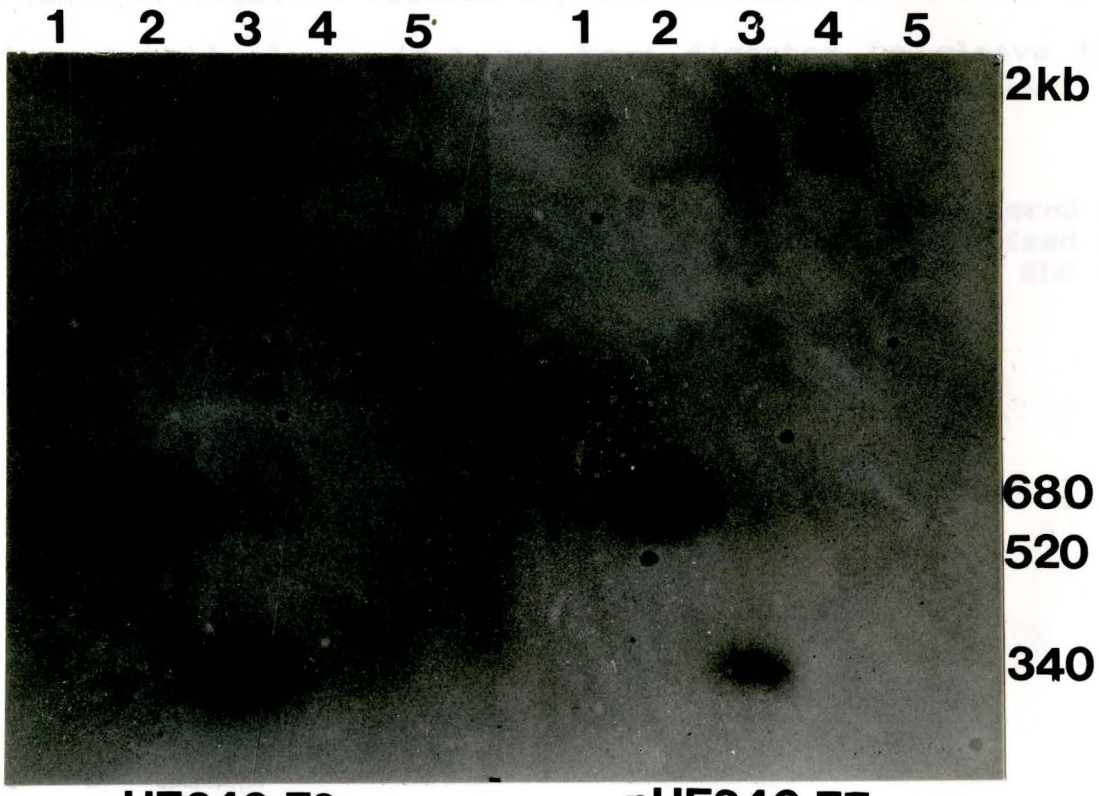


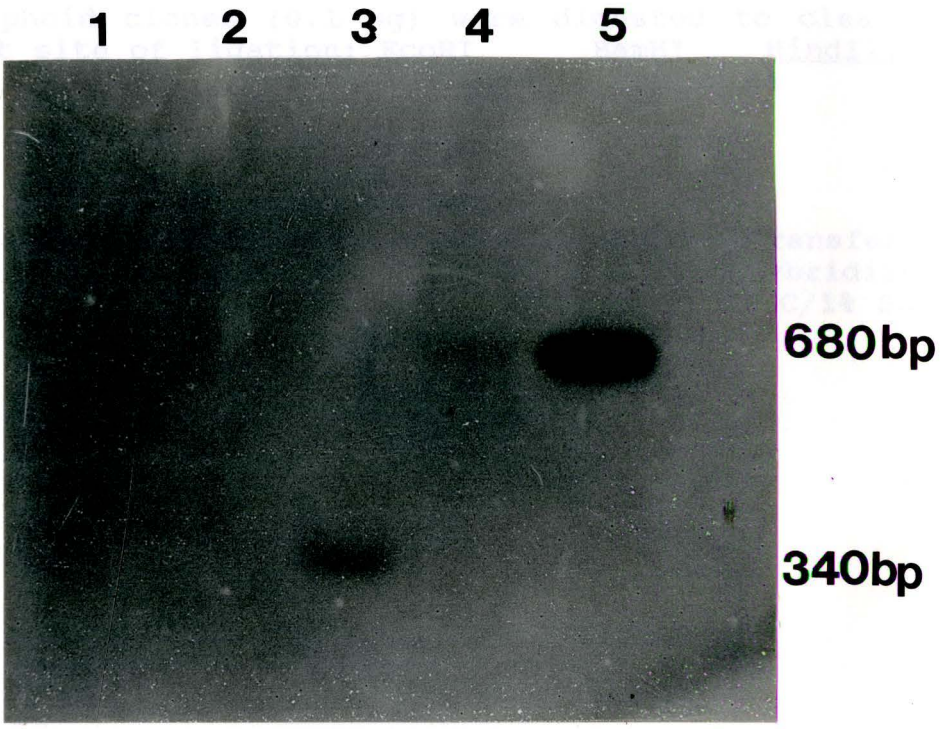
Figure 8.

Determination of stringency strengths in the elimination of cross-hybridization between alphoid families and subfamilies

Five alphoid clones (0.1 ug) were digested to cleave the insert at site of ligation: EcoRI BamHI HindIII
pHE340-64 pXBR-1 PHH550-31
pHE340-76
pHE340-77
Fragments were resolved on a 1% agarose gel and transfered to Gene Screen Plus membranes. The membranes were hybridized as indicated. Hybridization at 37 C, wash at 0.1xSSC/1% SDS 65 C.

1=pHH550-31
2=pXBR-1
3=pHE340-64
4=pHE340-77
5=pHE340-76

Fig 8



pHE340-76

Figure 9.

Determination of stringency strengths in the elimination of cross-hybridization between alphoid families and subfamilies

Seven alphoid clones (0.1 ug) were digested to cleave the insert at site of ligation: EcoRI BamHI HindIII

pHE340-9 pXBR-1 PHH550-31
pHE340-30
pHE340-64
pHE340-76
pHE340-77

Fragments were resolved on a 1% agarose gel and transfered to Gene Screen Plus membranes. The membranes were hybridized as indicated. Hybridization at 50 C, wash at 0.1xSSC/1% SDS 65 C.

1=pHE340-64
2=pHE340-9
3=pHE340-30
4=pHE340-76
5=pHE340-77
6=pHH550-31
7=pXBR-1

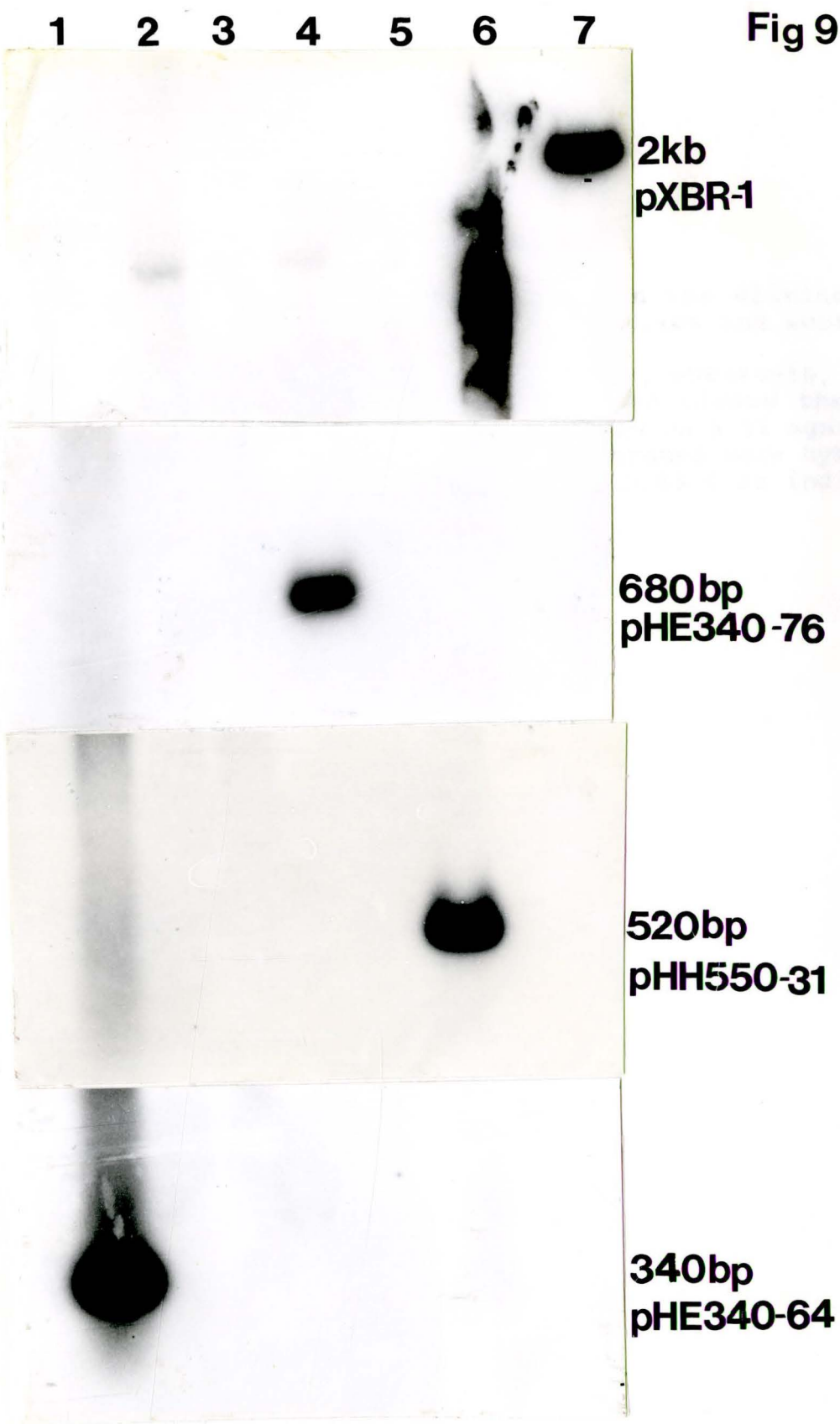

1
2
3
4
5
6
7
2kb
pXBR-1
680bp
pHE340-76
520bp
pHH550-31
340bp
pHE340-64

Fig 9

Figure 10.

Determination of stringency strengths in the elimination of cross-hybridization between alphoid families and subfamilies

Three remaining alphoid clones (pHE340-9, pHE340-30, pHE340-77) were digested (0.1 ug) with EcoRI to cleave the insert from the plasmid. Fragments were resolved on a 1% agarose gel and transfered to Gene Screen Plus. Membranes were hybridized at 50 C and washed at 0.1 x SSC / 1% SDS 65 C as indicated.

1=pHE340-64
2=pHE340-9
3=pHE340-30
4=pHE340-76
5=pHE340-77
6=pHH550-31
7=pXBR-1

Fig 10

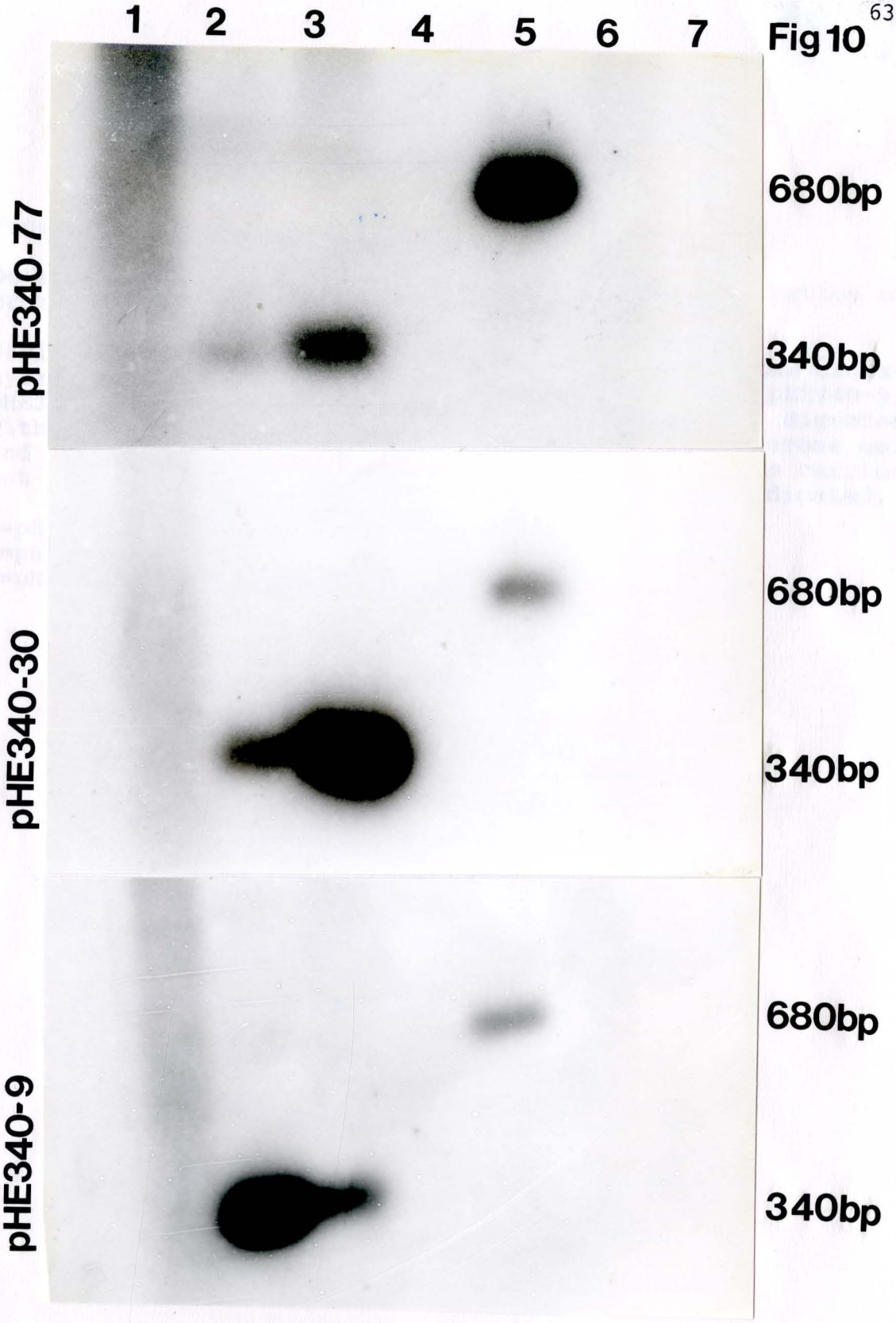

1
2
3
4
5
6
7
pHE340-77
680bp
340bp
pHE340-30
680bp
340bp
pHE340-9
680bp
340bp

Figure 11.

Determination of stringency strengths in the elimination of cross-hybridization between subfamilies

Stringency conditions were increased to eliminate the cross-hybridization between the EcoRI generated alphoids (pHE#40-9, pHE340-30, pHE340-77). 0.1 ug of each plasmid was digested with EcoRI. Fragments were fractionated on a 1% agarose gel and transfered to Gene Screen Plus. Hybridization was carried out at 55 C, wash at 0.04 x SSC / 1% SDS 65 C as indicated.

1=pHE340-9
2=pHE340-30
3=pHE340-77

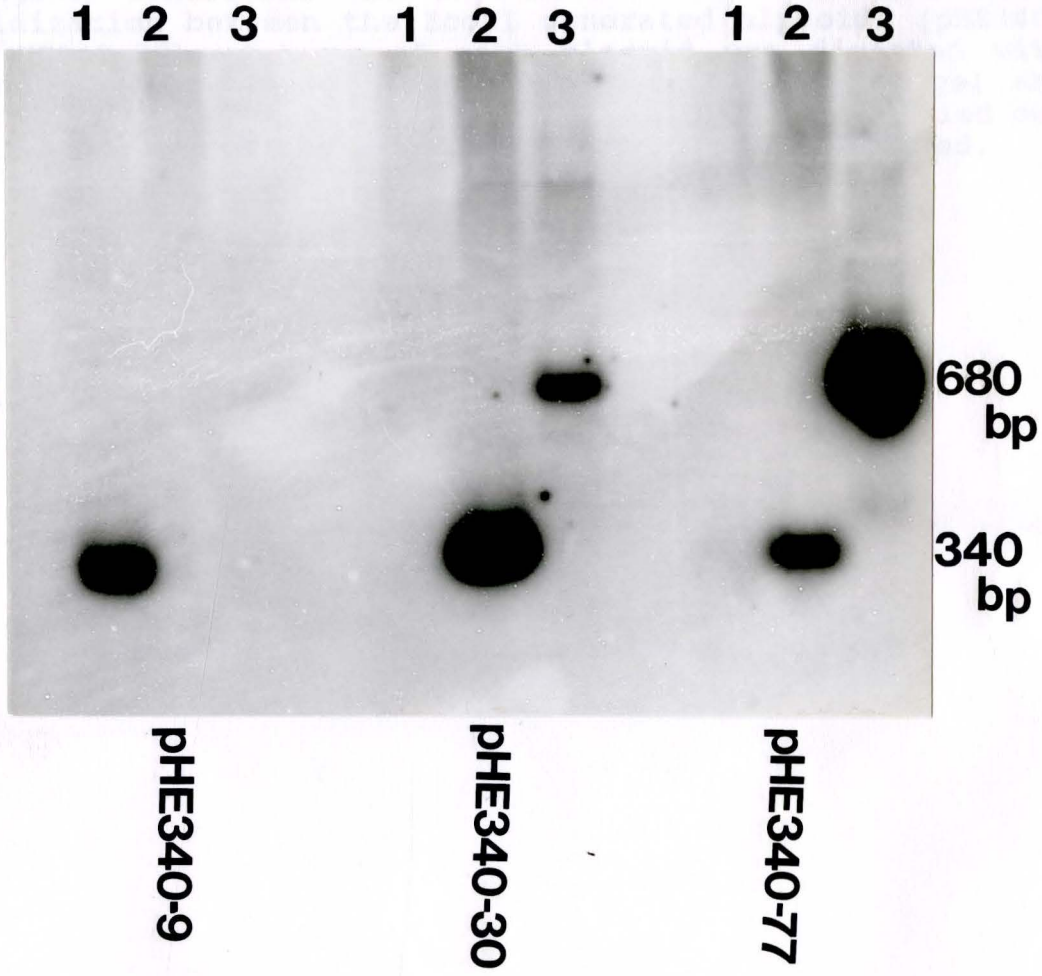

1 2 3 1 2 3 1 2 3
680 bp
340 bp
pHE340-9
pHE340-30
pHE340-77

Figure 12.

Determination of stringency strengths in the elimination of cross-hybridization between alphoid subfamilies

Stringency conditions were increased to eliminate the cross-hybridization between the EcoRI generated alphoids (pHE340-30, pHE340-77). 0.1 ug of each plasmid was digested with EcoRI. Fragments were fractionated on a 1% agarose gel and transfered to Gene Screen Plus. Hybridization was carried out at 55 C, wash at 0.002 x SSC / 1% SDS 65 C as indicated.

1=pHE340-77
2=pHE340-30
3=pHE340-9

Fig12

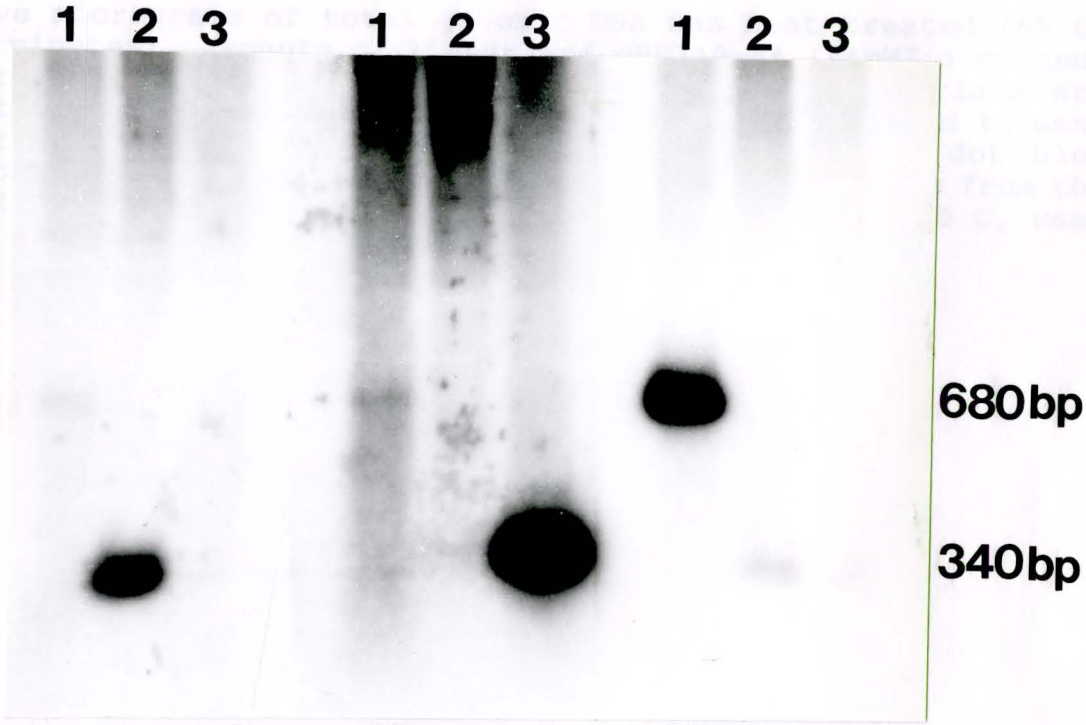


Figure 13.

Estimation of copy number of 340 bp human alphoid (pHE340-64) in human genome

Five micrograms of total genomic DNA was heat treated (65 C, 30 minutes). Amounts of linearized pHE340-64 (BamHI digested) were calculated to generate calibration (see Materials and Methods). Both genomic and plasmid DNA were tranfered to Gene Screen Plus with Schleicher and Schuell Minifold dot blot apparatus. The membrane was hybridized to insert only from the pHE340-64 clone. Hybridization was carried out at 50 C, wash at 0.1 x SSC / 1% SDS 65 C.

Fig 13

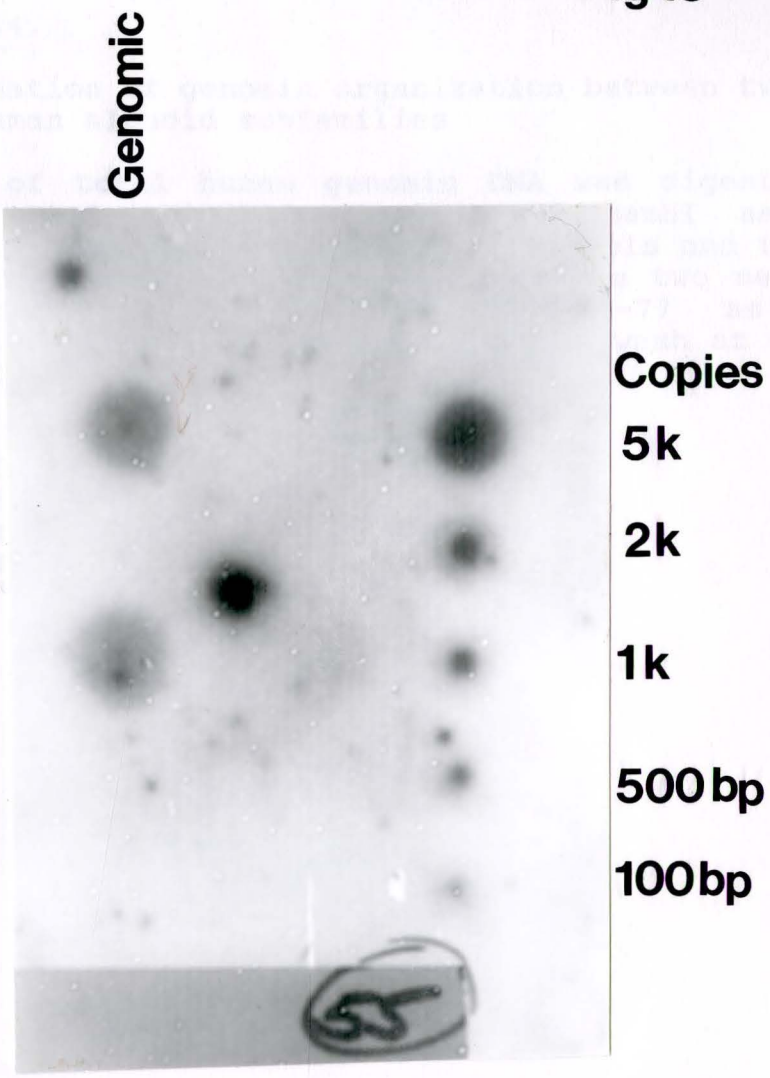


Figure 14.

Determination of genomic organization between two 680 bp EcoRI human alphoid subfamilies

0.5 ug of total human genomic DNA was digested with the enzymes AluI, HindIII, EcoRI, and BamHI as indicated. Fragments were resolved on 1% agarose gels and transfered to Gene Screen Plus Membranes. Each of the two membranes were hybridized to pHE340-76 or pHE340-77 as indicated. Hybridization was carried out at 37 C, wash at 2 x SSC / 1% SDS 65 C.

1 = AluI
2 = HindIII
3 = EcoRI
4 = BamHI

Fig14

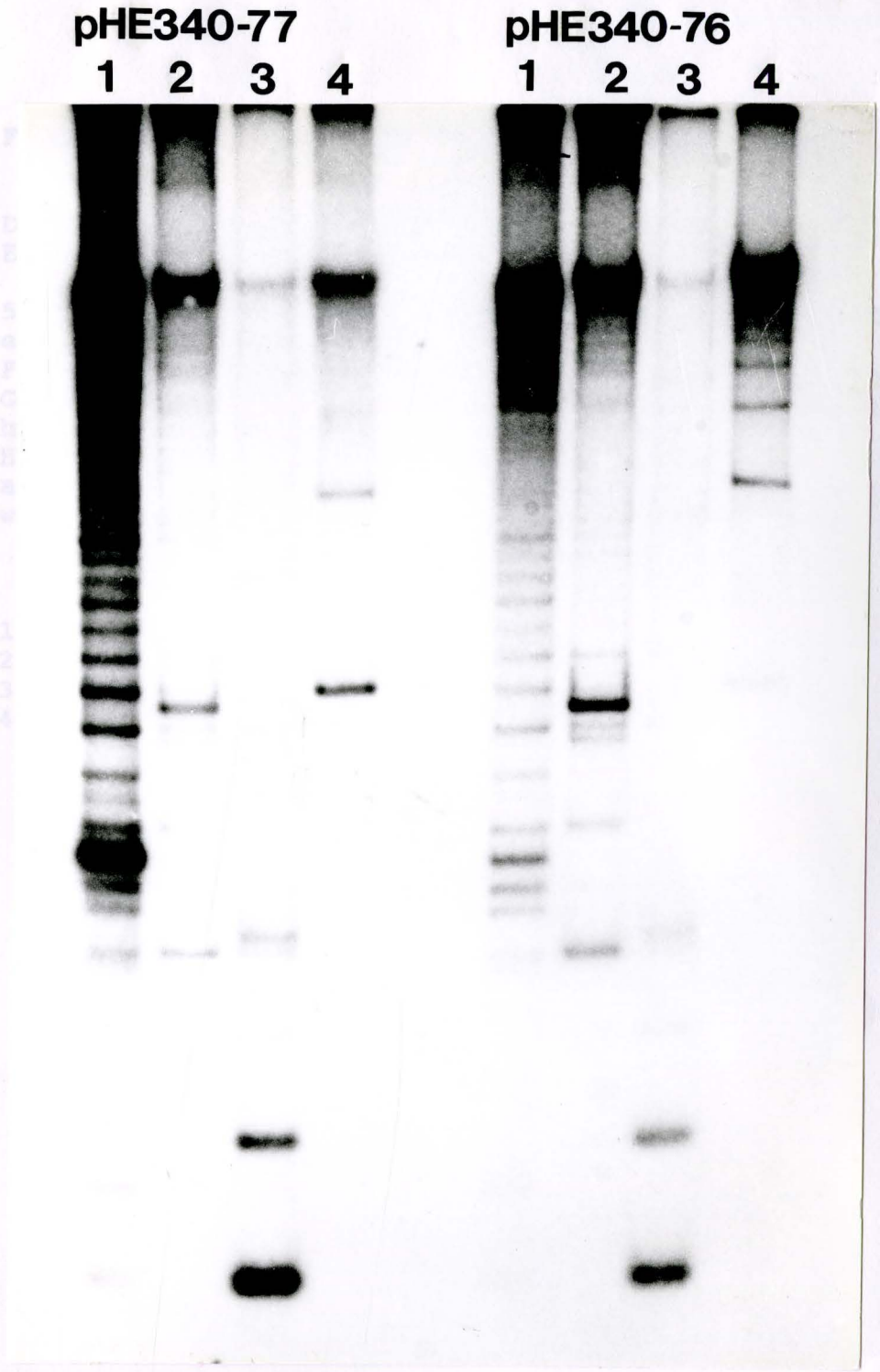


Figure 15.

Determination of genomic organization between two 680 bp EcoRI human alphoid subfamilies

5.0 ug of total human genomic DNA was digested with the enzymes AluI, HindIII, EcoRI, and BamHI as indicated. Fragments were resolved on 1% agarose gels and transfered to Gene Screen Plus Membranes. Each of the two membranes were hybridized to pHE340-76 or pHE340-77 as indicated. Hybridization of the pHE340-76 was carried out at 50 C, wash at 0.1 x SSC / 1% SDS 65 C. Hybridization of the p he340-77 was carried out at 55 C, wash at 0.002 x SSC / 1% SDS 65 C

1 = BamHI
2 = EcoRI
3 = HindII
4 = AluI

Fig15

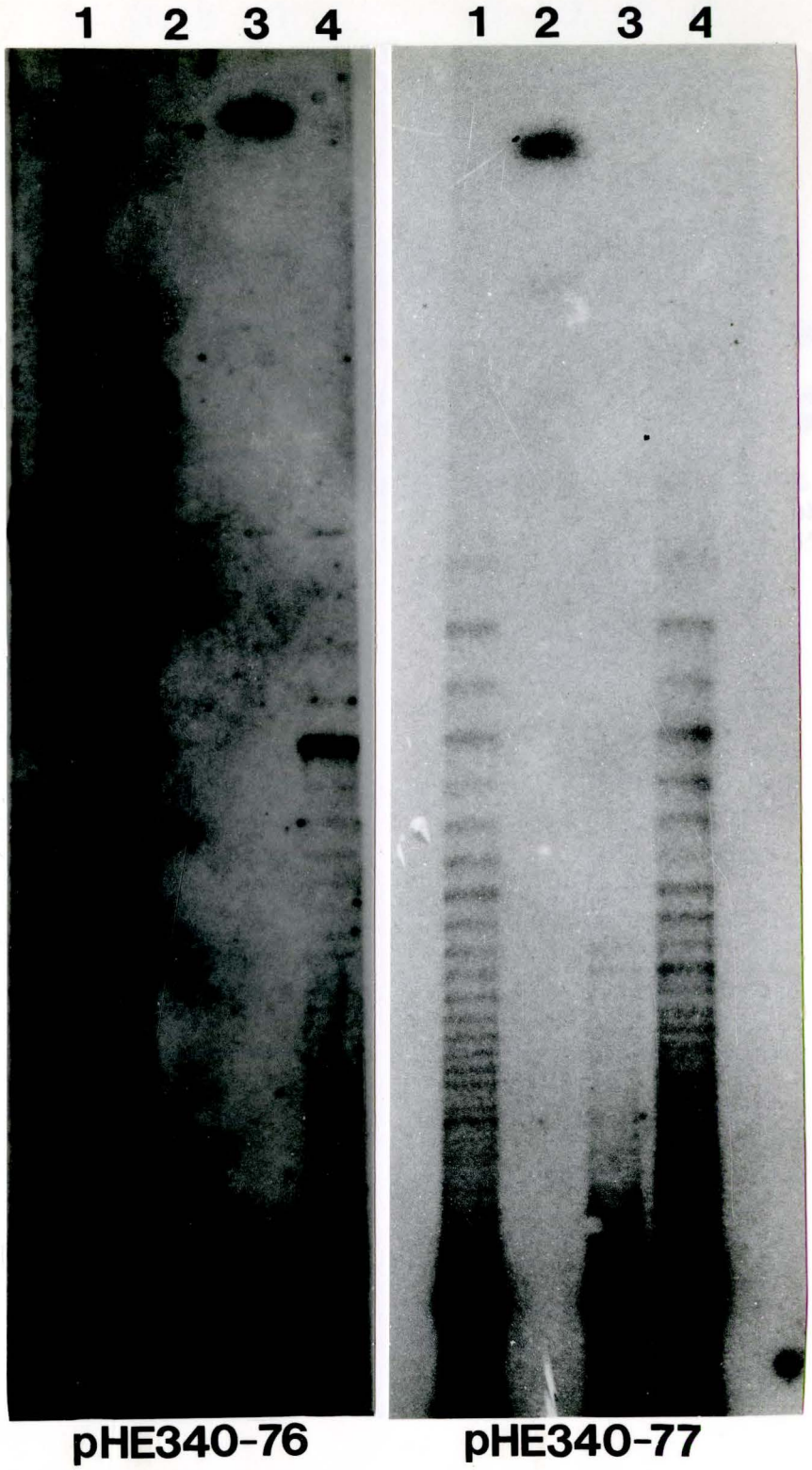


Figure 16.

Determination of genomic organization between two 340 bp EcoRI human alphoid subfamilies

5.0 ug of total human genomic DNA was digested with the enzymes AluI, HindIII, EcoRI, and BamHI as indicated. Fragments were resolved on 1% agarose gels and transfered to Gene Screen Plus Membranes. Each of the two membranes were hybridized to pHE340-9 or pHE340-64 as indicated. Hybridization for the pHE340-64 was carried out at 50 C, wash at 0.1 x SSC / 1% SDS 65 C. Hybridization of the pHE340-9 was carried out at 55 C, wash at 0.04 xSSC / 1% SDS 65 C.

1 = BamHI
2 = EcoRI
3 = HindII
4 = AluI

Fig16

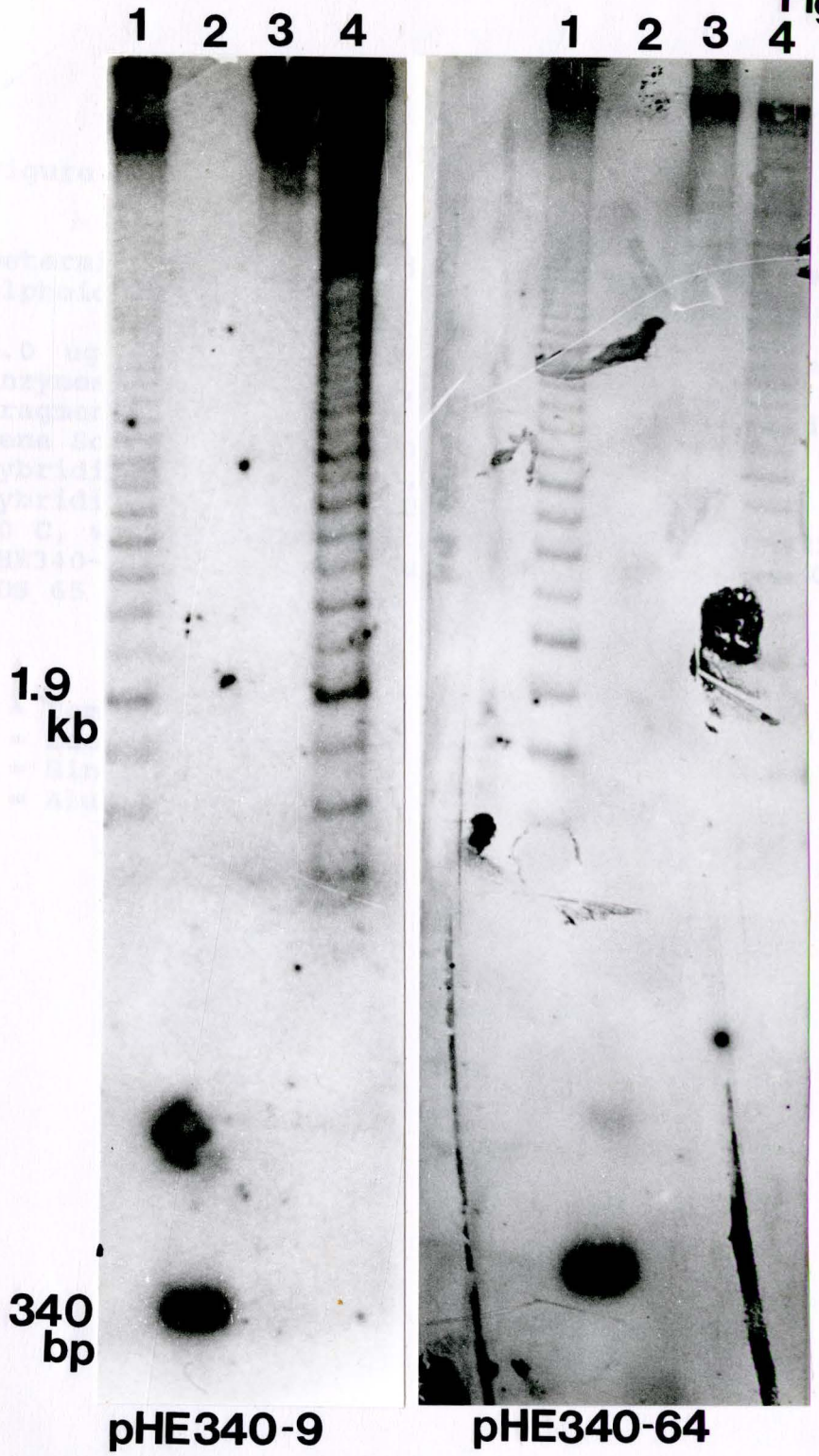


Figure 17.

Determination of genomic organization between three human alphoid families

5.0 ug of total human genomic DNA was digested with the enzymes AluI, HindIII, EcoRI, and BamHI as indicated. Fragments were resolved on 1% agarose gels and transfered to Gene Screen Plus Membranes. Each of the three membranes were hybridized to pHH550-31, pHE340-77, or pXBR-1 as indicated. Hybridization of the pHH550-31 and pXBR-1 were carried out at 50 C, wash at 0.1 x SSC / 1% SDS 65 C. Hybridization of the pHE340-77 was carried out at 55 C, wash at 0.002 x SSC / 1% SDS 65 C.

1 = BamHI
2 = EcoRI
3 = HindII
4 = AluI

Fig17

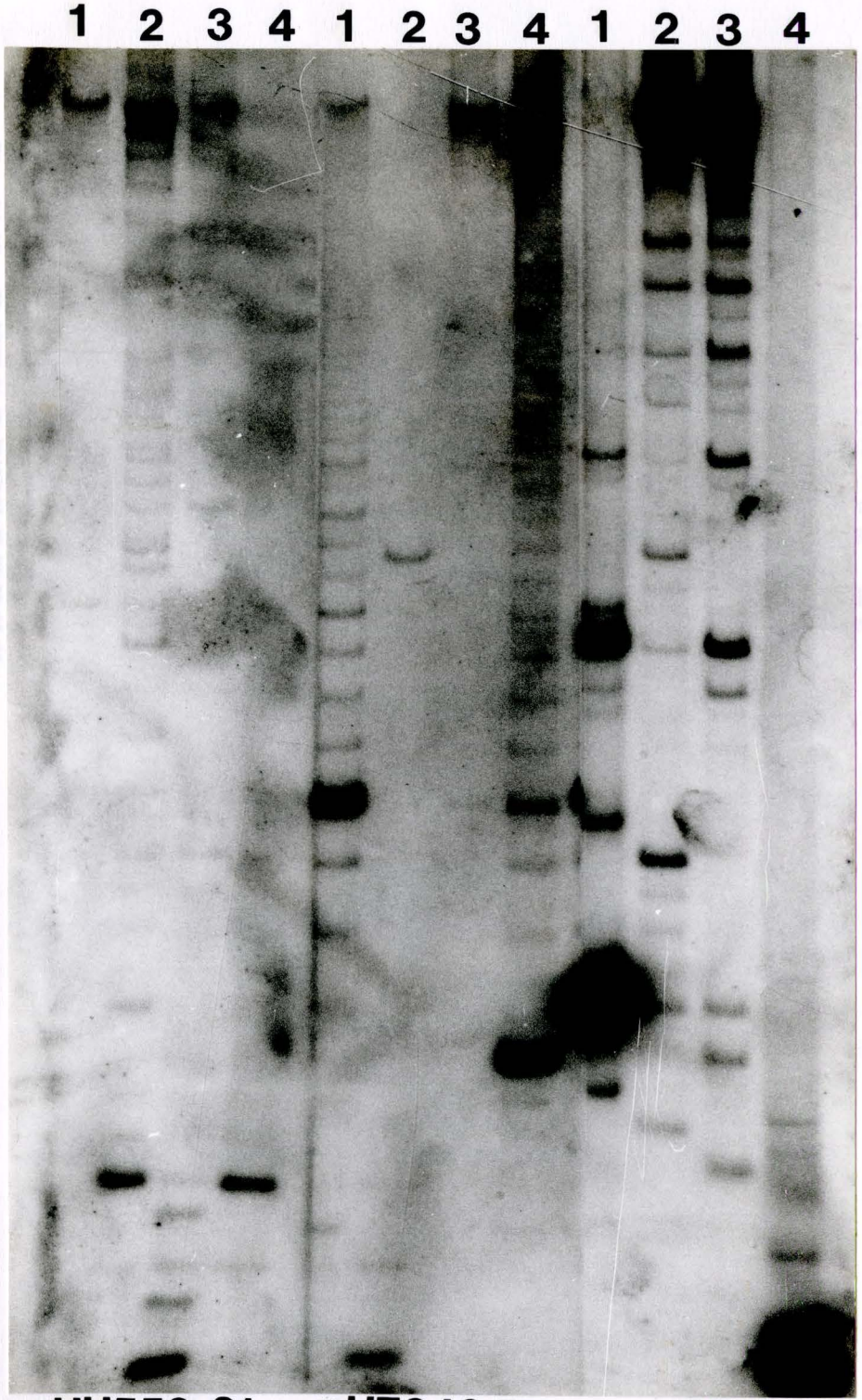


Figure 18.

Differences in retention between the plasmids pRM47 and pRM47340-1 in the yeast S. cerevisiae

Two freshly transformed strains of yeast, each with one of the plasmids (pRM47 or pRM47340-1) are grown in non-selective liquid media. Aliquots were taken from the media at several hour intervals and plated on both non-selective and selective (uracil deficient) plates. Non-selective plates contained all essential nutrients and indicated total number of cells present. Cells that survived on the selective media were uracil independent due to the the pRM47 plasmid. Graph represents plasmid retaining cells as URA +. A percentage of plasmid retaining cells was calculated as URA+ cells over total number of cells present. All aliquots plated were compared to the initial plating (t_0) to detect relative rate of plasmid loss.

Fig18

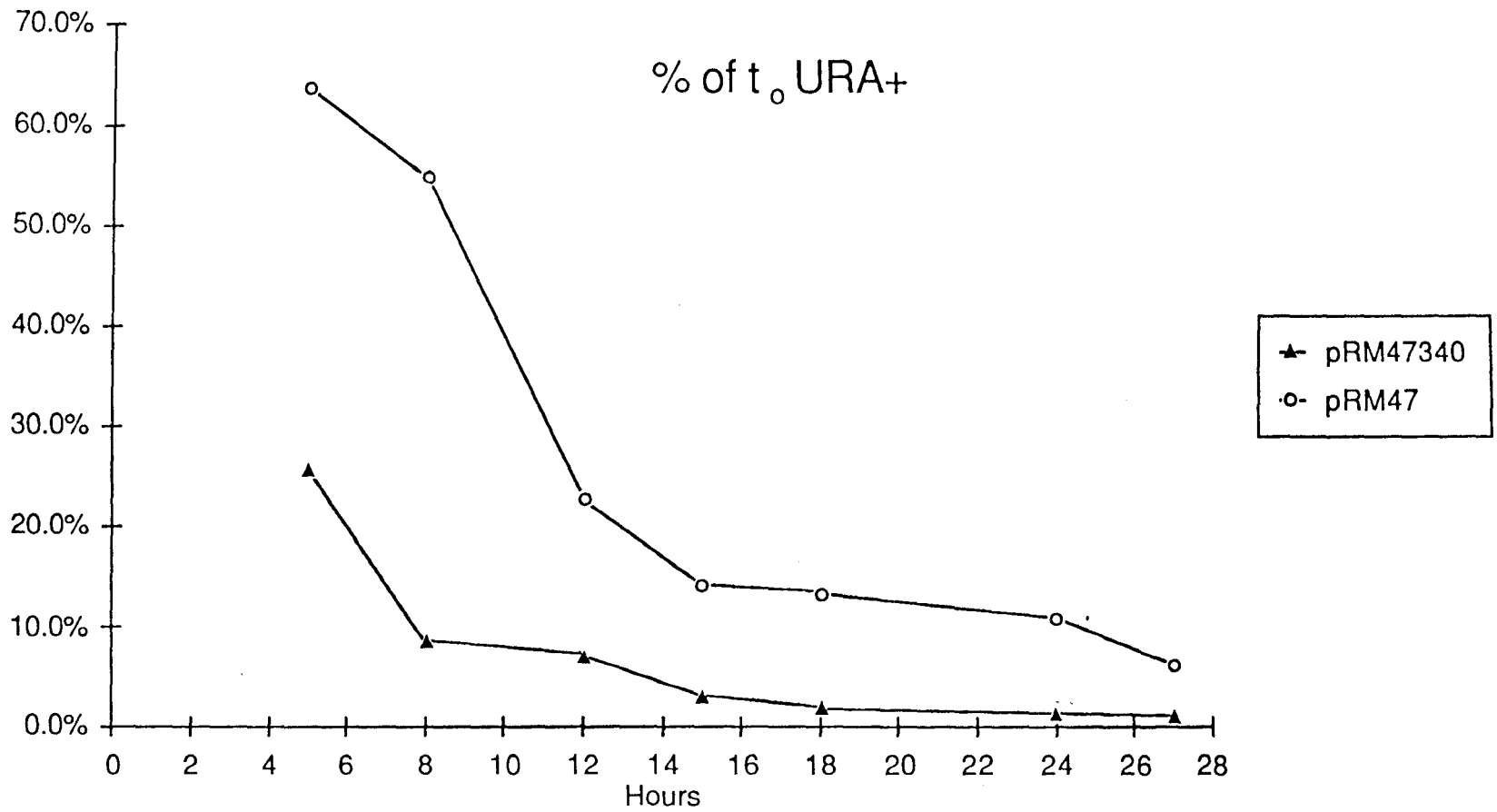
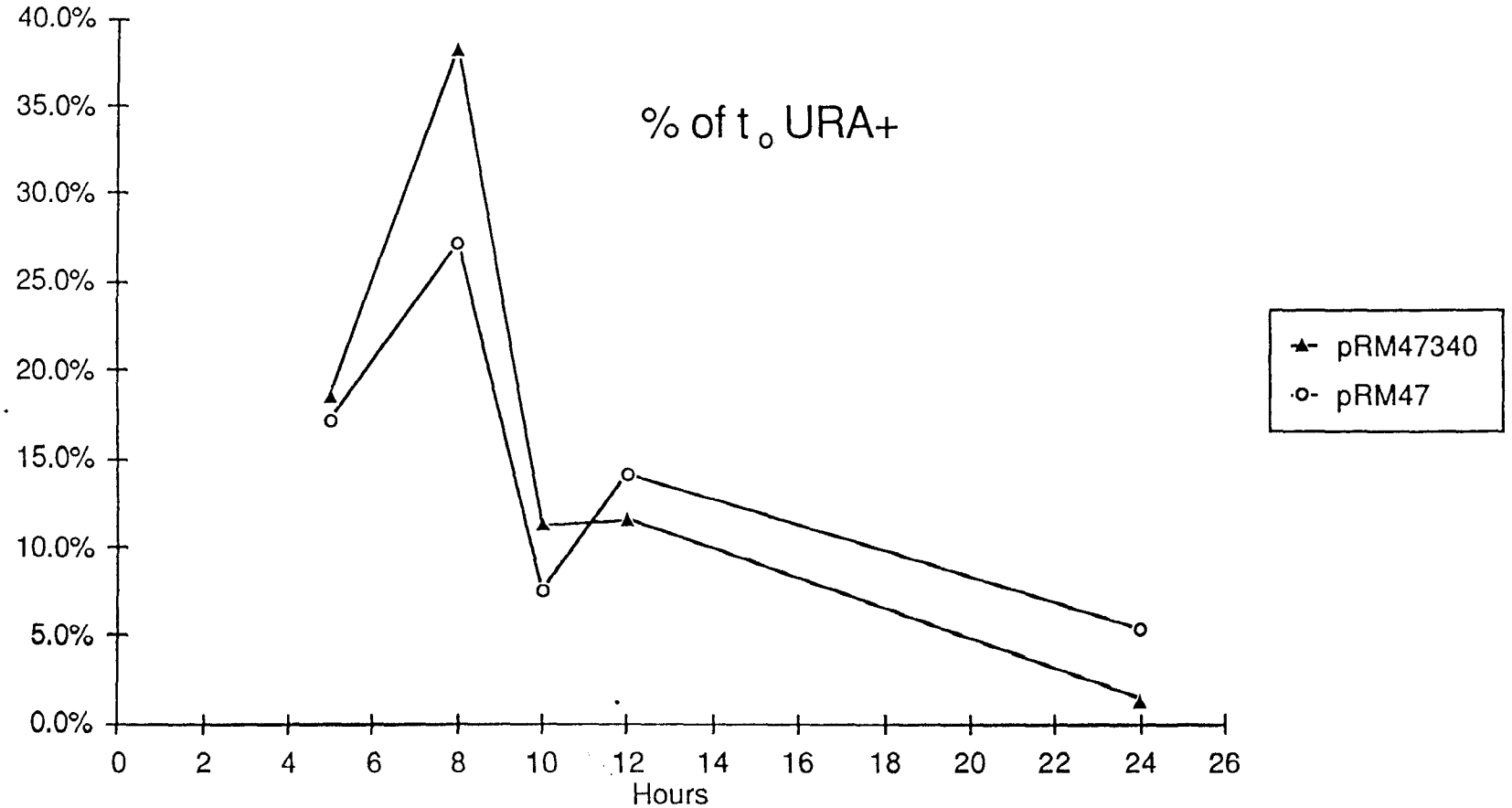

% of t o URA+
70.0%
60.0%
50.0%
40.0%
30.0%
20.0%
10.0%
0.0%
0
2
4
6
8
10
12
14
16
18
20
22
24
26
28
Hours
pRM47340
pRM47

Figure 19.

Differences in retention between the plasmids pRM47 and pRM47340-1 in the yeast S. cerevisiae

Two transformed strains of yeast, each with one of the plasmids (pRM47 or pRM47340-1) were taken from a previous functional assay and were grown in non-selective liquid media. Aliquots were taken from the media at several hour intervals and plated on both non-selective and selective (uracil deficient) plates. Non-selective plates contained all essential nutrients and indicated total number of cells present. Cells that survived on the selective media derived uracil component from the pRM47 plasmid. Graph represents plasmid retaining cells as URA +. A percentage of plasmid retaining cells was calculated as URA+ cells over total number of cells present. All aliquots plated were compared to the initial plating (t_0) to detect relative rate of plasmid loss.

Fig19


40.0%
35.0%
30.0%
25.0%
20.0%
15.0%
10.0%
5.0%
0.0%
% of t o URA+
0
2
4
6
8
10
12
14
16
18
20
22
24
26
Hours
pRM47340
pRM47

BIBLIOGRAPHY

Biro PA, Carr-Brown A, Southern EM, and Walker PMB, (1975). Partial sequence analysis of mouse satellite DNA, evidence for short range periodicities. J. Mol.Biol. 94: 71-86.

Bloom KS, and Carbon J, (1982). Yeast centromere DNA is in a unique and highly ordered structure in chromosomes and small circular minichromosomes. Cell 29:305 -317

Brennicke A, & Hemleben V, (1983). Sequence analysis of the cloned cucumis-melo highly repetitive DNA. Z. Natureforsch. 38c:1062-1065

Brutlag DL, (1980). Molecular arrangement and evolututionof heterochromatic DNA. Ann. Rev. Genetics 14:121-144

Clarke L, Hanspeter A, Fishel B, and Carbon J, (1986). Analysis of centromeric DNA in the fission yeast Schizosaccharomyces pombe. Proc. Natl. Acad. Sci. USA 83:8253-8257

Comings DE, and Okada TA, (1971). Fine structure of kinetochore in Indian Muntjac. Exp. Cell Res. 67:97-110

Corneo G, Ginelli E, and Polli E, (1967). A satellite DNA isolated from human tissue. J. Mol. Biol. 23:619-622

Corneo G, Ginelli E, and Polli E, (1970). Repeated sequences in human DNA. J. Mol. Biol. 48:319-327

Corneo G, Zardi L, Polli E, (1972). Elution of human satellite DNAs on a methylated albumin kieselghur chromatographic column: isolation of satellite IV. Biochim Biophys Acta 269:201-204

Davidson EH, Hough BR, Klein WH,(1975). Sctructural genes adjacent to interspersed repetitive DNA sequences. Cell 4(3): 217-238

Doering JL, Burket AE, Hanlon KM, and Schlegel DS, (1986). New subfamilies of human alphoid repetitive DNA. J. cell Biol. 103:491a

Doering JL, and Burket AE, (1985). Two new families of Human repetitive DNA. J. Cell Biol. 101:73a

Fitzgerald-Hayes M, Clarke L, Carbon J, (1982). Nucleotide sequence comparisons and functional analysis of yeast centromere DNAs. Cell 29:235-244

Furlong NB, Koenraad M, Flook B, and White J, (1986). Characteristic of site variation among clone of the 340-base pair tandemly repeated EcoRI family of human DNA. Biochemical Genetics 24:71-78

Grellet F, Delcasso D, Panabieres F, and Delseny M, (1986). Organization and evolution of a higher plant alphoid-like satellite DNA sequence. J. Mol. Biol. 187:495-507.

Grunstein M, and Hogness D, (1975). Colony hybridization: A method for the isolation of cloned DNAs that contain a specific gene. Proc. Natl. Acad. Sci. 72:3961.

Heiter P, Pridmore D, Hegemann JH, Thomas M, Davis RW, and Philippsen P, (1985). Functional selection and analysis of yeast centromeric DNA. Cell 42:913-921

Kurnit DM, and Maio JJ, (1973). Subnuclear redistribution of DNA species in confluent and growing mammalian cells. Chromosoma 42:23-36.

Jabs EW, Meyers DA, and Bias WB, (1986). Linkage studies of polymorphic, repeated DNA sequences in centromeric regions of human chromosomes. Am. J.Hum. Genet. 38-297-308.

Jokelainen PT, (1967). The ultrastructure and spatial organization of the metaphase kinetochore in rat cells. J Ultrastruct. Res. 19:19-44.

Jorgensen AL, Bostock CJ, and Leth Bak A, (1986). Chromosome-specific subfamilies within human alphoid repetitive DNA. J. Mol. Biol. 187:185-196.

Lohr D, Van Holde KE (1975). Yeast chromatin subunit structure. Science 188:165-166.

Maio JJ, (1971). DNA strand reassociation and poly-ribonucleotide binding in the African green monkey. J. Mol. Biol. 56:579-595.

Maio JJ, Brown FL, and Musich PR, (1981). Towards a molecular paleontology of primate genomes. Chromosoma 83:103-125.

Manuelidis L, (1976). Repeating restriction fragments of human DNA. Nucleic Acids Res 3:3063.

Manuelidis L, (1978).Chromosomal localization of complex and simple repeated human DNA. Chromosoma 66:23-32

McCutchan T, Hsu H, Thayer RE, and Singer MF (1982). Organization of African green monkey DNA at junctions between alpha-satellite and other DNA sequences. J. Mol. Biol. 157:195-211.

Mitchell AR, Beauchamp RS, Bostock CJ, (1979) A study of sequence homologies in four satellite DNAs of man. J. Mol. Biol. 135:127-149.

Mitchell AR, Gosden JR, and Miller DA, (1985). A cloned sequence p82H of the alphoid repeated DNA family found at the centromeres of all human chromosomesChromosoma 92:369-377

Pardue ML, and Gall JG, (1970). Chromosomal localization of mouse satellite DNA. Science 168:1356-1358.

Peacock WJ, Dennis ES, Rhoades MM and Pryor AJ, (1981). Highly repeated DNA sequence limited to knob heterochromatin in maize. Proc. Nat. Acad. Sci. USA 78:4490-4494.

Peterson JB, and Ris H, (1976). Electron microscopic study of the spindle and chromosome movement in the yeast *S. cerevisiae*. J. Cell Sci. 22:219-242.

Potter SS, and Jones RS, (1983). Unusual domains of human alphoid satellite DNA with contiguous non-satellite sequences: sequence analysis of a junction region. Nucleic acids res. 11:3137-3153.

Rigby P, Dieckmann M, Rhodes C, and Berg P, (1977). Labeling of DNA to high specific activity *in vitro* by nick translation with DNA polymerase I. J. Mol. Biol. 113:237-252.

Ris H, and Witt PL, (1981). Structure of the mammalian kinetochore. Chromosoma 82:153-170.

Roos UP, (1973). Light and electron microscopy of rat kangaroo cells in mitosis II. Kinetochore structure and function. Chromosoma 41:195-220.

Rubin CM, Deininger PL, Houck CM, and Schmid CW, (1980). Partial nucleotide sequence of the 300 nucleotide interspersed repeated human DNA sequences. Nature 284: 372-374.

Schmid CW, and Jelnik WR, (1982). The Alu family of dispersed repetitive sequences. Science 216:1065-1070.

Sharp PA, (1983). Conversion of RNA to DNA in mammals: Alu-like elements and pseudogenes. Nature 301:471-472.

Shoomkler Reis RJ, Srivastsava A, Beranek DT, Goldstein S, (1985). Human alphoid family of tandemly repeated DNA: Sequence of clonedtetrameric fragments and analysis of familial divergence. J. Mol. Biol. 186:31-41.

Singer MF, Highly repeated sequences in mammalian genomes. Int. Rev. Cytol. 76:67-112.

Smith CT, and Brown WRA, (1987). Structure of the major block of alphoid satellite DNA on the human Y chromosome. J. Mol. Biol. 195:457-470.

Southern EM (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 98:503-517.

Stinchcomb DT, Struhl K, and Davis RW, (1979). Isolation and characterization of a yeast chromosomal sequence. Nature 39-43.

Strauss F, and Varshavsky A, (1984). A protein binds to a satellite DNA repeat at three specific sites that would be brought into mutual proximity by DNA folding in the nucleosome. Cell 37:889.

Walbot V, and Goldberg RB, (1979). Plant genome organization and its relationship to classical plant genetics. In Nucleic acids in plants, editors TC Hall, JW Davies. Boca Raton, Fla. : CRC Press, c1979.

Weintraub H, (1980). Recognition of specific DNA sequences in eucaryotic chromosomes. Nucl. Acids Res. 8: 4745-4753.

Whitney FR, and Furano AV, (1983). The independent evolution of two closely related satellite DNA elements in rats. Nucl. Acids Res. 11(2): 291-304.

Willard HF, (1985). Chromosome-specific organization of human alpha satellite DNA. Am. J. Hum. Gen. 37:524-532.

Wolfe J, Darling SM, Erickson RP, Craig IW,Buckle VJ, Rigby PWJ, Willard HF, Goodfellow PN, (1985). Isolation and characterization of an alphoid centromere repeat family from the human Y chromosome. J. Mol. Biol. 182:477-485.

Wu JC, and Manuelidis L, (1980). Sequence definition and organization of a human repeated DNA . J. Mol. Biol. 142:363.

Yang TP, Hansen SK, Oishi KK, Ryder DA, Hamkalo BA (1982). Characterization of a cloned repetitive DNA sequence concentrated on the human X chromosome. Proc. Natl. Acad. Sci. USA. 79:6593-6597.